A New Travel Demand Model for Outdoor Recreation Trips

Xihe Jiao

Department of Architecture
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Queens' College

June 2018

Acknowledgement

I have been fortunate to receive support from many people in many ways through the course of this PhD, whose names may not all be enumerated. Their contributions are sincerely appreciated and gratefully acknowledged.

First and foremost, I want to offer my deepest gratitude and appreciation to my supervisor Dr. Ying Jin. His breadth and depth of knowledge and diligence have been and will always be an inspiration for me.

I wish to specially thank my advisor Prof. Philip James at the University of Salford, without whom this dissertation would not be possible. I thank him for his unwavering support and invaluable advice throughout. I also would like to thank his student and colleagues: Andrea Perz, Claire Gely and Paul Oldfield, who provided me valuable data which has been an important part of this research.

I wish to thank my colleagues in the Applied Urban Modelling team, with whom I have often had inspirational discussions. I am particular indebted to Steve Denman, who has patiently helped me to tackle numerous problems on GIS and data collecting

This research would not have been possible without the generous funding from China Scholarship Council. and I am grateful for the Department of Architecture and Queens' College for the support throughout my PhD.

Last but definitely not the least, I would like to give my deepest love to my families who have been always and unconditionally supporting me. This is particular in memory of my father, who deceased just before this research could be finished. Without him I could not become the person I am today, there will always be a special place in my life for him.

Thank you very much, everyone!

Abstract

Travel to outdoor recreational spaces belongs to a general class of research questions for understanding destination and travel mode choices. In travel demand modelling, discrete choice models (DCMs) have been applied to understand and predict a wide range of choices, such as how people choose among alternative destinations for jobs, homes, shopping, personal services etc. Surprisingly, DCMs have rarely been used to understand and model travel to outdoor recreational spaces. In the current literature for modelling travel to outdoor recreational spaces, the established models are Negative Binomial Regression (NBR) models, such as what was used in the UK NEA studies. However, these NBR models were developed to assess the effects of travel to outdoor recreational spaces at a national level, and they are not intended for assessing choices of individual sites. One reason for this is, as identified by previous studies, is that compared with the DCMs, the NBR models have certain limits on estimating people's choice behaviours. There is, therefore, no existing model that can represent and predict how people choose to travel to outdoor recreational spaces. Given the importance of outdoor recreational activities to urban land use planning and public health, this is a clear gap in the field.

The aim of this study is to develop a new travel demand model capable of representing and predicting travel to individual outdoor recreational sites. This is achieved by answering four main research questions: First, how to build the new model for outdoor recreational travel? Secondly, is the estimation accurate enough? Thirdly, to what extent can the new model be transferred to destinations outside the case study area? And, finally, how can city planners and designers use this new method? The new model draws upon ideas from random utility theory that underlies the conventional travel demand models to represent trip generation, trip distribution and mode choice. This research follows the standard modelling procedure: data collection and preliminary analysis, model calibration, model validation and model application. The data are collated from a wide range of sources that, importantly for model transferability, cover all areas in England. The new model has been calibrated for a case study area which spanned 14 selected districts in the North-West region. Validation of the new model is based on estimating the

numbers of trips to two outdoor recreational sites (Wigg Island and Wigan Flashes) and to nine English National Parks where data on visitor trips exist. In the final stage of the research, the new model is applied to estimate the changes that would arise from planning and design interventions in existing (Wigg Island and Moore Nature Reserve) and proposed (Arpley Country Park) sites.

At the end of this process, it is possible to show that the new model can predict the number of trips to individual destinations and that the model can be transferred to other outdoor recreation sites. Furthermore, the new model presented here is capable of predicting the changes in the volume and catchment of visits to an existing green space after land use planning or urban ecological interventions. This is a completely new theoretical model that is focused on understanding and quantifying the travel choices to outdoor recreation sites, which can inform decision makers by forecasting changes in outdoor recreational travel demand, according to different planning scenarios.

Content

List of Figures

List of Tables

Chapter 1 Introduction

In the field of transport modelling, it is common to develop computer simulation models that incorporate relevant theories and methods from more than one discipline to understand the behaviour of people and transport systems (Ortúzar & Willumsen, 2011). For instance, for journeys to work, schools and other destinations that are prone to traffic congestion (which has been the primary focus of transport modelling so far), the computer simulation models have incorporated engineering theories regarding traffic flows, the consumer choice theories regarding trade-offs people make among destination and travel options, and the effects of land use and urban design in shaping the destination and travel choices. Surprisingly, there are no interdisciplinary models of this type regarding the choice of outdoor recreation activities.

Travel to outdoor recreational spaces has increased significantly in the last decade. According to a report from the Natural England (2016), the total number of outdoor recreational trips has risen from 2.86 billion (2009-2010) to 3.12 billion (2015-2016) in England, generating more than 20 billion pounds in expenditures. Hence, it is not a surprise that there is a growing interest in understanding the outdoor recreational travel behaviour of residents. Natural England has funded the DEFRA (Department for Environment, Food & Rural Affairs) and Forestry commission to conduct a survey, called Monitor of Engagement with the Nature Environment (MENE) since 2009. This survey has provided robust evidence for the study of travel demand for outdoor recreational trips. A trip forecasting model built by Sen et al. (2012, 2014) is the only model based on the MENE data. However, it was developed for applications at the national and regional scale. Travel demand of individuals for outdoor recreation purpose remains poorly understood. The aim of this research is to build a new travel demand model for outdoor recreation activities, based on the mainstream theories and methods in the transport planning area. The new model can be used to estimate travel demand for any outdoor recreational destination. The model results can assist planners in assessing interventions of land use and landscape regarding their effects on outdoor recreation activities and associated benefits.

This kind of new travel demand model for outdoor recreation trips is essential because, relative to other primary disciplines that shape land use and the landscape, such as traffic engineering and estate finance, the effects of specific land use planning or landscape design interventions that aim to improve outdoor recreation activities remain poorly quantified. This puts those who wish to promote such projects at a disadvantage when debating short-term funding priorities and longer-term management of outdoor recreation spaces in the context of land use planning and landscape design.

The literature review for this research shows that one particular weak link that leads to this disadvantage is the poor understanding of how people actually travel to outdoor recreation destinations. This appears to be a field of research work that has fallen through a long-standing gap between transport planning and environmental studies. On the one hand, transport planning has over the years created a sophisticated understanding of and prediction models for practically every kind of travel except outdoor recreation travel. Although the benefits of outdoor recreation travel are believed to be significant, how people travel to outdoor recreation destinations remains poorly understood (Phaneuf & Smith, 2005). On the other hand, environmental studies, particularly environmental economics, have started to examine outdoor recreation travel, such as the UK National Ecosystem Assessment (UK NEA) studies in the UK using the MENE survey mentioned above (Sen et al., 2012, 2014). However, so far, such work has only focused on national scale effects and is currently unable to examine interventions at a local scale (Sen, 2015, private communication).

Although a national scale picture is essential, local interventions such as the planning and design of a new country park or the improvements of existing green space are crucial for planning, designing and creating outdoor recreation destinations. Nowadays, practically all the other planning decisions are made based on substantial evidence regarding how people access land use sites for jobs, homes, shopping, personal services, logistical depots, etc. (Boyce & Williams, 2015). This asymmetry in knowledge results in individuals leaning on a biased evidence base that cannot properly value the benefits of outdoor recreation space relative to interests in traffic engineering or estate finance.

The aim of this dissertation is to develop a new travel demand model that can represent and predict travel to individual outdoor recreational sites. The new model draws upon ideas from mainstream transport modelling that underlies transport and land use planning in representing how frequently people travel, where they choose to go, and what means of transport they adopt. The model links the geographical distribution of visits to key land uses, landscape design and urban design features at a local level, such as the distribution of population among neighbourhoods, the location of recreational sites, transport accessibility and environmental characteristics, to the outcomes of travel decisions. The resulting quantification of the impacts of policy interventions is expected to make a significant improvement to the empirical basis for decisions on investment, regulation, and planning of outdoor recreation sites. More specifically, this research aims to address four main research questions: First, how to build the new model for outdoor recreational travel? Secondly, is the estimation accurate enough? Thirdly, to what extent can the new model be transferred to destinations outside the case study area? And finally, how can city planners and designers use this new method?

The rest of the dissertation is organised as follows. In Chapter 2, the existing literature in related fields is reviewed, starting from travel demand modelling in general and then focusing on discretion choice models (DCMs) which underlies much of travel forecasting modelling. In the latter part of Chapter 2, another branch of travel forecasting modes, the Negative Binomial Models (NBR) is reviewed, per its application in the outdoor recreation studies, such as those applied in the UK NEA studies.

The overall structure of the analytical process of model building and model application is then presented in Chapter 3, which acts as a roadmap to what is presented in Chapters 4 and 5. This chapter shows how model calibration, validation, application and various case studies will fit together.

Chapter 4 presents model calibration and validation. It starts examining the data collected from a wide range of sources, covering all areas in England to facilitate model transferability. The socio-demographic profiles of the residents at the origin of the outdoor recreation trips and the environmental attributes of the destination sites are collated from a diverse range of sources and used jointly, for the first time, as explanatory variables in the new model. The new model has been calibrated for a case study area which spans selected districts in the North-West region. The validation of the new model is based on estimating the numbers of trips to two outdoor recreational sites—Wigg Island Nature Reserve and Wigan Flashes—and to the ten English National Parks. For model validation, the results are then compared with data either collected through on-site observations or extracted from the annual reports produced by the National Parks' management team. These data have not been used in model calibration.

In Chapter 5, the new model is applied to evaluate the implications of alternative patterns of future developments in the Upper Mersey Estuary (UME) area. Three scenarios are tested through the new model: the 'business as usual' scenario with no dramatic changes on the site; the 'Development boom' scenario in which the central focus is economic growth; and the 'Nature is the Key' scenario, where environment conservation is the main focus. These scenarios were developed by Dr Andrea Drewitt of the Ecosystems and Environment Research Centre at the University of Salford. The new model was applied to estimate the changes of visits to both existing sites (i.e., Wigg Island and Moore Nature Reserve) and a new proposed outdoor recreational site (Arpley Country Park), in line with the different planning strategies.

Finally, Chapter 6 concludes this research by assessing the strengths and weaknesses of the new model and, in this light, suggests directions for future work.

Chapter 2 Literature Review

2.1 Overview

Besides the age-old belief that outdoor activities are good for the body and spirit, in recent years, there has been a growing evidence base showing that outdoor recreation is closely associated with human health and wellbeing (Bateman et al., 2014; Fuller et al., 2007; Tzoulas et al. 2007). Proximity to urban green space is associated with increased levels of physical activity (Booth et al., 2000; Humple et al., 2004). Outdoor recreation contributes to the prevention and management of conditions such as coronary heart disease, diabetes and obesity, costly expenses for the National Health Service (Liu et al., 2007; Richardson, 2013). People are happier where they have accessible green space, and individuals have lower mental distress, improved attention and lowered blood pressure (Hartig et al., 2003; Ottosson & Grahn, 2005; White, 2013). Accessing to green spaces in an urban area is also believed to improve levels of community cohesion and promote social inclusion (Kim & Kaplan, 2004).

This evidence appears to have started to influence how people perceive the benefits of outdoor recreation. For instance, in the UK, health and exercise have become among the most frequently cited motivations for visiting the outdoors. For example, in a systematic survey on Managing the Engagement with the Natural Environment (MENE) that has been going for seven years, the proportion of outdoor recreation visits where health and exercise were cited as a motivation rose from 34 per cent in 2009 to 47 per cent in 2016 (Natural England, 2017).

Given the importance of the health and wellbeing of citizens, this growing evidence base should exert a greater influence on how the outdoor recreational spaces are planned and designed. In the field of transport planning, it is common practice nowadays to develop computer simulation models that incorporate relevant theories and methods from more than one discipline to understand the behaviour of people and transport systems (Ortúzar & Willumsen, 2011). This chapter contains an examination of such relevant theories and methods regarding the systematic analysis and modelling of outdoor recreational trips. The relevant literature suggests that the analysis and modelling of outdoor recreational trips appear to fall between two poles: on the one hand, the four-step method of travel modelling, first initiated in the 1950s and then developed significantly during last 50 years to cover all types of travel (Boyce & Williams, 2015), has rarely touched outdoor recreational travel; on the other hand, the regression models that have been developed and applied within the environmental geography and economics for predicting outdoor recreational travel have not engaged the field of travel demand modelling; instead, such models place a unique emphasis on the use of land cover and land use data which is absent from the four-step models (Perino et al., 2011; Sen et al., 2014).

In order to establish a general framework of definitions and model structures, this chapter first reviews the conventional four-step models in Section 2.2, the core travel choice theories in Section 2.3, and the application of such models in Section 2.4. Similarly, there is a review of regression models, particularly the Poisson regression and its variation, the Negative Binomial Regression (NBR) in Section 2.5, and the application of such models in Section 2.6. The last Section (2.7) summarises the findings from this chapter and offers four specific research questions in light of the literature review.

2.2 The structure of conventional travel forecasting modelling

Conventional transportation modelling is built upon four-step travel models which weres initially formed in the 1960s. The broad structure of this approach has remained more or less unchanged despite significant improvements in the modelling technique since then (Boyce & Williams, 2015). The standard structure includes four steps: first, trip generation to calculate the number of trips originating from each residential area for travel by residents (and similarly workers from areas of employment); secondly, trip distribution to estimate the number of trips per respective destinations; thirdly, modal choice to estimate travellers' choices among the means of travel; the final step is traffic assignment to calculate the flow of vehicles and people, as appropriate for each of the means of transport in their respective transport networks.

Since the research here is concerned with the generation, distribution and mode choice of outdoor recreational travel demand rather than the traffic on the networks, the first three steps are the most relevant, and they will be reviewed below. The conventional trip assignment modelling is, quite understandably for transport planning purposes, focused on the rush hours of the day when the capacity of the transport networks is under highest pressure. For studying outdoor recreational travel, this fourth step is not particularly relevant, mainly because the outdoor recreation journeys are usually made outside the rush hours and on transport networks that are not subject to the most intense congestion.

2.2.1 Trip generation

Trip generation aims at predicting the number of trips generated by each land parcel (i.e., zone) of a study area. The earliest method is called the growth-factor method developed by Fratar (1954). It estimates future trips by multiplying the current number of trips by a growth factor, which typically consists of variables such as population, income and car ownership. This method is easy to understand, but using a growth factor to estimate the number of trips means simplifying the complex travel systems crudely (Ortúzar & Willumsen, 2011). Attempts to improve these estimates have been made by calculating a total number of trips by a linear regression by zone and socioeconomic characteristics of the household. This is a more advanced statistical method compared with the growth factor method. However, there are problems with this technique. Firstly, zone models can only explain the variation in trips made between zones; hence, a major challenge are those main differences happening on a personal trip on an inter-zone level. Secondly, zone-based regression is conditioned by the nature and size of the zone (Douglas & Lewis, 1970). Modellers have tried to reduce inter-zone variations by using a smaller analysis unit. For example, at the beginning of the 1970s, the zone was first replaced by the household. Then, the question turned to linear regression itself: what if an independent variable exerts a non-linearity influence on the total number of trips?

From the late 1960s, an alternative appeared and quickly became the preferred trip generation method in the UK: category analysis (Wooton & Pick, 1967). The method is based on the assumption that trip generation rates are relatively stable over time for classes of household; thus, the number of trips per household by given category can be estimated as a function of household characteristics. Then, in the late 1970s, the issue came back to the question as to whether the household is a small enough unit to capture the variations between individuls. Therefore, a trip maker-based trip generation model was developed by Supernal et al. (1983) based on the household category model and has since become popular.

The advantages of the personal categorised model include the sample size required to prepare a person category model, which can be several times smaller than that needed to estimate a household-category model. Secondly, demographic changes can be more easily accounted for and forecasted in a person-category model. The primary function of person category model can be written as:

$$T_i = N_i \sum_j a_{ji} t_j \qquad \text{Equation 1}$$

Where T_i denotes the total number of trips made by the individuals of zone *i* (all categories together), N_i is the number of inhabitants of zone *i*, and a_{ji} is the percentage of inhabitants of zone i belonging to category j (the category could be any demographic or socio-economic group, such as sex, age, car ownership etc., which is subject to empirical tests). t_j is the number of trips made during a certain time period by a homogenous population in category *j*.

One weakness of this method is that changes to accessibility are assumed to have no effects on trip generation. This assumption is made because there are few extant data to show how changes in accessibility affect trip generation. Daly (1997) proposed a frequency choice logit model to incorporate the effects of accessibility. It is a logit form model which will be reviewed later, estimating the total number of trips by first calculating the probability that each individual would choose to make a trip. The utility of making a trip is usually specified as a linear function of explanatory variables such as income, car ownership, and household size, as well as accessibility. Where there is suitable data, this new component can enable the outcome of a travel assignment step to feed back into the trip generation step. In practice, as the current observations on trip generation are limited, a standard approach is to assume that the number of outdoor recreational trips per person per year remains unaffected by traffic conditions. Analysis from the MENE data, for instance, adopts this approach.

2.2.2 Trip distribution

Following the estimation of the number of trips from each zone, the next step is the modelling of the number of trips attracted to each destination's zone. As with the trip generation function, the earliest method used is the growth factor method. This process is heavily reliant on the accuracy of the base year trip matrix, and it is not capable of accounting for either travel cost changes or the cells in the trip matrix unobserved in the base year (Ortúzar & Willumsen, 2011).

As an improved form of the growth factor model, the gravity model, which was developed initially from an analogy from physics, that is to say from with Newton's gravity law, became popular very quickly in the 1960s due to the accuracy of estimations the gravity model can make (Voorhees, 1955). A total number of trips attracted to a destination zone was estimated by the total number of trips from both ends (origin and destination) and generalised function of travel cost. However, the gravity model cannot make good sense of how it could be related to travel behaviour. Therefore, people made efforts to develop models from different angles. One of the most cited papers in the history of urban transportation and regional science was written by Wilson (1967), who developed the gravity model through information theory—the entropy maximising principle. Still, the primary principle of people's choice behaviour is believed to lay in the fields of economics and psychology. The entropy maximising principle was not thought to be consistent with the theories in those two areas until the 1980s (Miyagi, 1984). As a result, the accusation of the arbitrary use of irrelevant (physical) perceptions on transportation analysis also affected the use of entropy maximising models in the early days (Daly, 2013).

The essence of the research into the motivation of travel forecasting since the 1970s came from the fields of economics and psychology. The key researcher in this development was Daniel McFadden, who developed the discrete choice random utility maximising (RUM) framework in the 1970s (McFadden, 2000), the most robust transportation modelling method today. Also, as mentioned above, it was shown by Miyagi (1984) that entropy maximisation could be seen as, in principle, entirely consistent with the RUM concept. Therefore, all entropy work could then be reinterpreted as models of choice made by utility-maximising travellers. The basic RUM-based discrete choice model forms are the multinomial logit model (McFadden, 1978) and the nested logit model (Daly & Zachary, 1978; Williams, 1977), used in this research to study outdoor recreational travel demand. Discrete choice models will be reviewed in Section 2.3 in detail, for their importance in understanding travel choices.

2.2.3 Modal choice

Modal choice modelling in the early years, particular in Europe, was dominated by trip-end models. It means the number of people who traveled by car was estimated based on car ownership, residential density and distance from the central area. Local public transport trips were determined as a proportion of total person trips at both ends based on the observed use for public transportation. This method was criticised because it is insensitive to changes in public service (Stopher & Meyburg, 1975).

This was followed by the development of the entropy maximising framework which was used for distribution models; the gravity model has also been used in mode choice, and the development of modal choice modelling has then followed the path of development of distribution models. Since the 1970s, the DCMs have gradually dominated the modal choice modelling process. The primary challenge for the modal-split model since the 1990s has become finding a rigorous behavioural basis for the location of the modal split in the travel forecasting procedure, either before or after trip distribution (Boyce & Williams, 2015). Three significant studies contributed to the assessment of model structure: a systematic review of 24 multi-step models of travel demand models in UK practice by Bly et al. (2002), a critical review of the response parameters for different market segments and the relative position of destination and modal choices by MVA (2005). RAND Europe (Rohr, 2005) specifically investigated the structures and parameters for destination and mode choice. In conclusion, for commuting purposes, the combination of destination first and then mode was appropriate. For the primary, secondary and tertiary education, shopping, and other travel including recreation, destination choice should follow modal choice (Rohr, 2005).

From the reviews above, it is not difficult to tell that the RUM-based DCMs dominate two out of three transport modelling steps—trip distributions and modal choice. Therefore, in the next section, the focus is on discrete choice models in some detail.

2.3 Discrete choice models

As reviewed in the last section, since the 1970s, instead of arbitrarily using the physical perception of transportation analysis, analytical models of choice behaviour have been developed in the fields of economics and psychology, with a focus on greater precision and more sensitivity to the behaviour of individuals (Ben-Akiva, 1973; Bruton, 1975; Domencich & McFadden, 1975). Random utility maximisation (RUM)-based Discrete Choice Models (DCMs) have become the dominant method in travel forecasting research (McFadden, 2000).

The goal of random utility models is to understand the behavioural process that leads to an individual's choices (Train, 2009). The RUM postulates that individuals act rationally and possess perfect information when they make choices—that is to say, the individuals always select the option which maximises their personal utility (net benifit) subject to legal, social, physical and budgetary (both in time and money terms) constraints (Ortúzar & Willumsen, 2011). However, the RUM theory distinguishes two types of factors that influence an individual's choice behaviour: those factors that are observed by the modeller, and those factors that cannot be observed by the modeller. Such unobservable factors are assumed to follow specific probability distributions as random elements in the individual's perception of the utility of a travel destination or travel mode; it is this assumption that gives rise to the name *random utility theory*. This assumption then makes it possible to account for the observed choices and influencing factors in a rigorous model of choice behaviour that can be empirically calibrated.

Random utility models that are applied to the choices of discrete options (such as outdoor destinations or transport modes) are defined as discrete choice models (DCMs) calculating the probability of individuals choosing a given option as a function of their socioeconomic characteristics and the relative attractiveness of the alternative. The DCMs can be developed for either individuals or groups of individuals. The DCMs about individuals will require a massive amount of data which few existing data sources can yet provide. The DCMs that are based on sample surveys need to be structured as models for groups of individuals. In such cases, the DCMs are defined as 'aggregate' choice models which will not capture all the diversity and variability of choices, although the random utility model structure is still capable of accounting for the probabilistic distribution of the unobserved factors of influence (Train, 2009). The discrete choice models come in many forms. The most popular forms are the single-level multinomial logit model (McFadden, 1974), the nested logit model (McFadden, 1978) and mixed logit model (McFadden & Train, 2000). Each model form has its own strengths and analytical limits, which will be reviewed in the following sections.

2.4 Single-Level Multinomial Logit model

The main principle of random utility modelling is best explained through a simple, single-level discrete model with a multinomial utility function. Assuming that an individual labelled n faces J options, the model separates the utility U_{nj} from option j into two parts: one is obtained from known parameters, labelled V_{nj}; the other part ε_{nj} is unknown ($j = 1,2, \dots, J$). Thus,

$$U_{nj} = V_{nj} + \varepsilon_{nj}, \forall J \qquad \text{Equation 2}$$

where the equation assumes that each ε_{nj} is independent, $\forall J$ means it assumes to be true for all options in J. Among the many alternative assumptions that can be made of the probabilistic distribution of the error term, the most analytically convenient is to assume that it has a Gumbel distribution as:

$$f(\varepsilon_{nj}) = e^{-\varepsilon_{nj}} e^{-e^{-\varepsilon_{nj}}} \qquad \text{Equation 3}$$

with its cumulative distribution written as

$$F(\varepsilon_{nj}) = e^{-e^{-\varepsilon_{nj}}} \qquad \text{Equation 4}$$

Note that the Gumbel distribution is analytically very similar to the normal distribution, but leads to a much more tractable function form for the choice probabilities (McFadden, 1974). The assumption of the Gumbel distribution means that the probability that the individual n chooses alternative i is

$$P_{ni} = \int \left(\prod_{j\neq 1} e^{-e^{-(\varepsilon_{ni}+V_{ni}-V_{nj})}} \right) e^{-\varepsilon_{ni}} e^{-e^{-\varepsilon_{ni}}} d\,\varepsilon_{ni} \qquad \text{Equation 5}$$

Some algebraic manipulation of $\mathrm{P_{ni}} = \int \left(\prod_{\mathrm{j}\neq 1} \mathrm{e}^{-\mathrm{e}^{-(\varepsilon_{\mathrm{ni}}+\mathrm{V_{ni}}-\mathrm{V_{nj}})}} \right) \mathrm{e}^{-\varepsilon_{\mathrm{ni}}} \mathrm{e}^{-\mathrm{e}^{-\varepsilon_{\mathrm{ni}}}} \mathrm{d}\,\varepsilon_{\mathrm{ni}}$

Equation 5 results in a succinct, closed form expression known as a logit-based DCM:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} \qquad \text{Equation 6}$$

Where *e* is the exponential function, a multinomial utility function V_{nj} representing utility for individual *n* choosing alternative *i.* The most common utility function is linear in parameters $V_{ni} = \beta' x_{ni}$ where β' is the parameter to be estimated, and x_{ni} is a vector of the observed variables relating to the influences regarding individual *n's* choice of the alternative i. In other words, the choice probabilities become

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{ni}}} \qquad \text{Equation 7}$$

It is worth noting that the logit formula was derived from assumptions that, for any two alternatives, the relative odds of choosing one over the other are the same no matter what other options are available or what the attributes of the other alternatives are. This assumption is called the independence of irrelevant alternatives (IIA) property (Luce, 1959). McFadden (1976) later stated that the IIA property has two advantages and one limitation. The advantages are that the IIA property implies that model calibration can be carried out by studying conditional choice in a small subset of the full set of alternatives, and, secondly, it allows quick analysis of the effects of introducing a new option (McFadden, 1976). The IIA property is nevertheless implausible for alternative sets containing choices that are close substitutes, where two of the alternatives may be considered by individuals to be so similar that the utilities for one option significantly correlate with the other (such as bus services on the same route but run by different companies). In order to overcome this weakness, a more general model form is required.

2.4.1 Nested logit model

As discussed above, the researcher is often unable to capture all sources of correlation explicitly so that the unobserved portions of utility are correlated and IIA assumption is violated. The nested logit model, initially developed by Williams (1977) and Daly and Zachary (1978), is one of the most widely used model forms to mitigate the IIA limitation. The nested logit model belongs to a large class of models called generalised extreme value (GEV) models. The unifying attribute of these models is that the unobserved part of utility for all alternatives are jointly distributed as generalised extreme value (Train, 2009); this means this class of models allows for correlations over alternatives and, as it is a generalisation of univariates, extreme value distribution that is used for standard logit models. When all correlations are zero, the GEV distribution becomes the product of independent extreme values distributions, and the GEV model becomes a standard, single-level logit as presented above. Therefore, the hypothesis test on the correlations within a GEV model can be used to verify whether the standard logit provides an accurate representation of the substitution patterns.

In the nested logit model, the set of alternatives faced by an individual can be partitioned into subsets, called nests. IIA holds within each nest but not in general for options in different nests. The function was initially developed by McFadden in 1978 and then rewritten by Train (2009) into two standard logit models, which are much easier to understand. Train first decomposed the observed component of the utility U_{nj} (the utility of individual *n* will get from choosing alternative *j*) into two parts: let the set of alternatives *j* be partitioned into *K* non-overlapping subsets B_1, B_2, …, B_K and called nests, a part labelled W that is constant for all alternatives within a nest B_K, and a part labelled *Y* that varies over alternatives *j* within the nest, ε_{nj} is the part of the utility that cannot be observed:

$$U_{nj} = W_{nk} + Y_{nj} + \varepsilon_{nj} \qquad \text{Equation 8}$$

W_{nk} depends only on variables that describe nest B_K. These variables differ over nests but not over alternatives within each nest. Y_{nj} depends on variables that define alternative *j*. These variables vary over alternatives within nest *k*. With this decomposition of utility, the nested logit probability can be written as the product of two standard logit probabilities as shown in Equation 10 and Equation 12: P_{ni} is the probability that an individual *n* chooses alternative *i* is equal to, $P_{ni|B_k}$ the probability of individual n chooses the alternative *i* over all the options within nest B_K, multipled by the P_{nB_k}, the probability that the individual *n* chooses the nest B_K over all the nests:

$$P_{ni} = P_{ni|B_k} P_{nB_k} \qquad \text{Equation 9}$$

where

$$P_{ni|B_k} = \frac{e^{Y_{ni}/\lambda_k}}{\sum_{j \in B_k} e^{Y_{nj}/\lambda_k}} \qquad \text{Equation 10}$$

$$I_{nk} = ln \sum_{j \in B_k} e^{Y_{nj}/\lambda_k} \qquad \text{Equation 11}$$

$$P_{nB_k} = \frac{e^{W_{nk} + \lambda_k I_{nk}}}{\sum_{l=1}^{K} e^{W_{nl} + \lambda_l I_{nl}}} \qquad \text{Equation 12}$$

The parameter λ_k is a measure of the degree of independence in unobserved utility among the alternatives in nest k. A higher value of λ_k means greater independence and less correlation over different nests. I_{nk} is the log of dominator in $P_{ni|B_k}$, this value links different levels of nests, often called the inclusive value or inclusive utility first identified by Ben-Akiva (1973). This structure has been widely used in transport modelling, and the earliest nests structure was proposed by Williams (1998). The nests are drawn upon different combinations of frequency (F), destination (D), mode (M) and route (R) as shown in Figure 2.1.

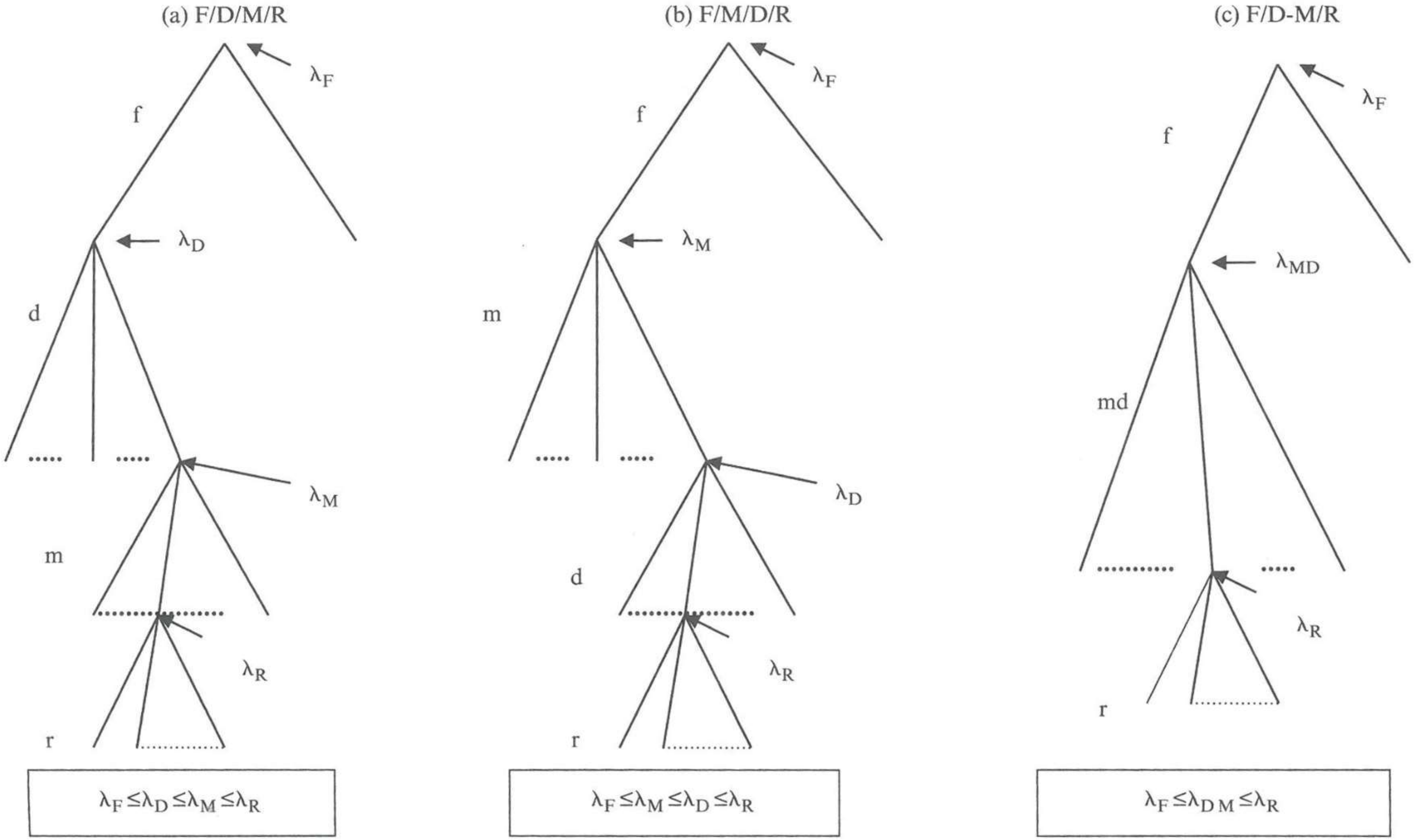


Figure 2.1. Tree structures for alternative nested logit models over multiple dimensions: frequencies (F), destinations (D), modes (M) and routes (R), wither parameter restrictions appropriate to the selected hierarchical structure. Source: Boyce and William (2015).

At any node and every level, the probability of choosing an option in the relevant set of choices would be represented by a standard multinomial logit model, with a value representing expected utility that an individual receives from the choices among the alternatives in the lower nest. When inputting the inclusive value (I_{nk} in Equation 11) into a standard multinomial logit model, a parameter λ_k is generated for level *k*, which determined the sensitivity of behavioural responses at different level *k* of a tree to changes in the generalised and composite costs. There is a restriction on parameters λ, wherein the parameters λ of the multinomial logit models associated with different level should not increase as the choices progress towards the top of the tree structure (Williams, 1977) because the choices at the bottom of the tree should be subject to the least random variations. The different structure could be subject to empirical test, and this criterion will be used to eliminate inappropriate model structures.

In summary, the nested logit model can be used to compare between choices that are modelled using logit-based DCMs, which in this research involves destination and mode choices. It can be used to answer the question: what is the best model structure regarding the order of modelling for destination and mode choice?

This leaves us with one other important issue to consider. Since the data sources in the case study are based on sample surveys (such as MENE), the new model is likely to be defined as aggregate choice models, which cannot fully capture the diversity and variability of individual behaviour. However, by defining the parameters of the utility functions as statistical distributions rather than constants, it will be able to test the variability among individuals. This is done through another alternative to the single-level multinomial logit model, which is called the mixed logit model and it is reviewed below.

2.4.2 Mixed Logit Model

The mixed logit model is defined by the functional form of its choice probabilities. The most straightforward derivation, and most widely used in recent applications, is based on random coefficients (McFadden & Train, 2000). The parameters of the deterministic proportion of the model are replaced with random coefficients that incorporate individual heterogeneity, and the random component of each parameter will be shared across choice alternatives (Phaneuf & Smith, 2005). $L_{ni}(\beta_n)$ is the probability of individual *n* choosing alternative *i* conditional on β_n which can be written exactly the same as the standard logit model:

$$L_{ni}(\beta_n) = \frac{e^{\beta_n' x_{ni}}}{\sum_j e^{\beta_n' x_{nj}}} \qquad \text{Equation 13}$$

The only difference between a random coefficients mixed logit model and the standard logit model is the coefficients β vary over individuals in the population with density $f(\beta)$. The researcher observes the x_{nj}, which is the vectors of observed variables but not β_n; if the researcher observed β_n then the choice probability would be standard logit. However, the researcher does not know β_n and therefore cannot condition on β; the unconditional choice probability P_{ni} is therefore the integral of Equation 13 over all possible variables of β_n

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right) f(\beta) d\beta \qquad \text{Equation 14}$$

In most applications, $f(\beta)$has been specified to be normal or lognormal distributed (Revelt & Train, 1998). The lognormal distribution is useful when the coefficient is known to have the same sign for every individual, such as the coefficient of travel cost which is known to be negative for everyone when they make choices.

In other words, the mixed logit model allows for random taste variation and correlation among individuals (Train, 2009). Through analysing the spread of the utility parameter functions, the model will be able to pinpoint the X that may need to be refined in data collection, such that the diversities and variabilities among the individuals can be better analysed.

The model also has its own limitations. When using the mixed logit model to make predictions, coefficients have to be simulated, which means the result of the correlation coefficient is a mean of numerous draws. Therefore, the mixed logit model is really only capable of testing variations of taste on the same variable among different individuals in the model calibration process, and it is a tool for exploring where in the future more data should be collected, in order to capture the diversities and variabilities of behaviour better as revealed by its findings.

Studies in the discrete choice random utility approach over the last two decades have been extensive, resulting in a broad application of increasingly sophisticated econometric models in transportation and other fields as market research, consumer behaviour, and environmental economics (Day, 2013). The next section will look at the development of travel demand modelling in particular, with focus on how DCMs have become the dominant method in the transportation field, and what the studies of DCMs have covered in the travel forecasting field, thus closing the section with a focus on the DCMs' application to the outdoor recreational trips.

2.5 Application of travel demand modelling

Transportation studies before the 1950s were designed to ameliorate street and highway congestion (Shuldiner, 2013). Tools such as desire lines were developed to measure traffic flow and used to identify traffic 'choke-point' and make plans to alleviate it. Driven by the increase of travel movement caused by the economic boom and in-migration from rural areas after the Second Word War in the industrial countries, there was a shift in transportation research from current to future orientation studies (Boyce & Williams, 2015), known as travel forecasting studies, to assist in the preparation for long-range plans for highways.

In order to understand how the RUM-based DCMs came to dominate travel demand modelling, one has to go back to the earliest innovation: the survey conducted by the state highway agencies. The first home interview survey regarding travel behaviour in some 125 urban areas was in 1953 (Lynch, 1959). The collected data included household characteristics as well as motor vehicle trips made by household members, and then the data were aggregated into census zones for analysis purpose. The Detroit Metropolitan Area Transportation Study (1955-1956) and the Chicago Area Transportation Study (1955-1959) are the earliest travel forecasting studies based on this data, and these studies have established the pattern for decades of transportation studies that followed. These studies aimed to explain the correlation between zone household characteristics and zone trip-making behaviours mainly for commuting trips. The limitations of research in this era (the 1960s – 1970s) are that they were exclusively focusing on land use (i.e., housing and work places), as the determinate of working trip demands. Variations between individual trip makers were only expected to be minimised by census zones, which grouped homogenous households on purpose (Shuldiner, 2013).

Such a concept was not sufficient to explain all travel movements. In particular, the techniques do not catch the competition between same land use within similar distance ranges. Mitchell and Rapkin (1954) were the first of the few researchers to raise the concept of the motivation of travel: they suggested that travel demand should be considered with changes related to activities or social structure, as well as land use and distance or travel time changes. Meanwhile, many industrial countries were struck by the first oil shock in the latter part of 1973. Large-scale capital projects gave way to transportation system management and a range of shorter-term planning schemes. The transportation modelling required analyses of higher precision and more sensitivity to the behaviour of individuals (Ben-Akiva, 1973; Bruton, 1975; Domencich & McFadden, 1975).

Innovations have been made since the 1970s which have significant theoretical and practical relevance today (Boyce & Williams, 2015). The distinct quality of travel demand models in this era is ‘disaggregate’ and ‘probabilistic’. These models were designed to capture the likely choice of an individual faced with a set of well-defined travel options. Various statistical techniques were applied to the formulation and estimation of these multivariate models: discriminant analysis by Waner (1962), Quarmby (1967) and McGillivray (1970), and probit analysis by Lisco (1967) and Lave (1969). Parameters of logit models were estimated by Warner (1962) and Stopher (1969). All of these led up to, as reviewed above, the development of Discrete Choice Models (DCMs) on the basis of Random Utility Maximisation (RUM) theory. In the following decade, various efforts were made to relax the IIA restrictions of the primary form of DCMs—the standard logit model. The most popular variations include the nested logit (NL) model by Williams (1977) and Daly and Zachary (1978), and a standard form for generalised extreme value (GEV) models by McFadden (1978), and the mixed logit model (ML) by Ben-Akiva and Bolduc (1996) and McFadden and Train (2000).

During the last two decades, the development of travel demand modelling was led by regulation innovations in the US (i.e., the Clean Air Act Amendments of 1990 and the Intermodal Surface Transportation Efficiency Act of 1991) and governments' initiatives in the UK (i.e., the New Labour Government elected in 1997). In particular, travel forecasting was required to be able to cover a broader range of policy proposals with greater focus on detail (Table *2.1*).

Table 2.1 *Areas Needing to be Included in Travel Forecasting Modelling*

Areas to be included in the transportation forecasting:
Land use policies
Highway projects
Heavy rail, modern bus, tram and light rail transport systems
Highway capacity reallocation
Central area parking and park-and-ride proposals
Road pricing schemes of varying description
Priority for 'slow modes' (bicycle and walk)
A range of policies to promote 'smart travel choices'

Source: Boyce and Williams (2015).

Several groups of travel forecasting systems have been developed with a great focus on details, heavily used in today's travel forcasting for transport planning. Two famous highway network analysis models are SATURN (simulation and assignment of traffic in urban road networks) (Hall et al., 1980) and CONTRAM (continuous traffic model) (Taylor, 2003). These methods for matrix estimation are derived from traffic accounts on links and junctions. The second group of models which grew dramatically in the last two decades are activity-based microsimulation models. The systems applied most widely over the past ten years are the TRANSIMS (Smith et al. 1995) in the United States, and PTV's VISSIM (Fridrich et al., 2000) and SIA's PARAMICS (Van Vuren, 2010) in the UK. The third group which was developed since the 1990s is the use of models in incremental form (i.e., incremental multinomial logit model and incremental nested logit model), in which the base matrix derived from observed data source serve as a reference from which growth and network changes are assessed. Early examples were the START (Roberts & Simmonds, 1997) and APRIL models (Bates et al., 1996).

All of these models focus on peak-time transportation capacity issues and the connections with supporting business activities; hence, the models are concerned with commuting and business trips (particularly at morning peak). Trips to other land uses are mainly focused on journeys to services (such as education, retail, personal business, etc.). Journeys to outdoor recreation spaces are rarely modelled because they are likely to take place outside the weekday rush hours and are not directly connected to productive or other business activities.

Although outdoor recreation trips are seldom mentioned in the transportation modelling, the studies of outdoor recreation in the environmental and economic fields are not rare. Research using discrete choice method focus on a single habitat/site for their own economic interests. For example, the fresh water or coastline recreations are among the most extensively studied areas (Table *2.2*). Parsons and Kealy (1992) looked at fresh-water recreation at Wisconsin Lakes through the nested logit model; Kling and Thomson (1996) investigated sport fishing in California using the two levels nested logit mode; Parsons, Plantinga, and Boyle (2000) studied fishing lakes in five sites in China's lakes region; Parsons, Massey, and Tomasi (2000) investigated beach recreation in Delaware, New Jersy, Maryland, and Virginia; and Hicks and Strand (2000) studied publicly accessible recreation beaches along the western shore of the Chesapeake Bay in Maryland.

A common weakness of such studies is the lack of transferability, as it makes the method difficult to apply for land use planning. One reason for the lack of general outdoor recreational activities studies is that there is virtually no data existing regarding general outdoor recreation before the MENE survey, which has only happened since 2009; the first data were published in 2010. A study carried out by Sen et al. (2013) as part of the UK National Ecosystem Assessment (NEA) project used a different modelling theory—the variation of the Poisson regression, the Negative Binomial Regression (NBR). The next section will review this theory and its application to the research of UK NEA.

Table 2.2 *Discrete Choice Modelling-based Outdoor Recreation Studies*

Author	Recreation type	Model
Parsons and Kealy (1992)	Fresh-water recreation at Wisconsin lakes	Nested logit
Feather (1994)	Fresh-water at Minnesota lakes	Standard logit
Shaw and Ozog (1999)	Five sites in Maine, three in Nova Scotia, New Brunswick, and Quebec, Canada	Nested logit
Kling and Thomson (1996)	Sports fishing in California	Nested logit
Parsons and Hauber (1998)	Recreational fishing in Maine	Nested logit
Parsons, Plantinga, and Boyle (2000)	Fishing lakes in Maine	Nested logit
Jones and Lupi (1997)	Recreational fishing in Maine	Nested logit
Parsons, Massey, and Tomasi (2000)	Beach recreation in Delaware, New Jersey, Maryland, and Virginia	Nested logit
Peters, Adamowicz, and Boxall (1995)	Fresh-water fishing in Southern Alberta, Canada	Standard logit
Hicks and Strand (2000)	Publicly accessible recreation beaches along the western shore of the Chesapeake Bay in Maryland	Standard logit

2.6 The Poisson regression and negative binomial regression

The Poisson regression is the basic count data model form, which is a limiting case of the negative binominal distribution (Cameron & Trivedi, 2013). The binominal distribution is related to a discrete random variable which is the number of successes in a sample of n observations during a time interval of given length. For example, in Equation 15, λ is the expected number of successful occurrences in the interval; the probability of success, which should be equal over all observations, is:

$$p = {}^{\lambda}/_{n} \qquad \text{Equation 15}$$

The probability of exactly y successes in the given time interval can be estimated by multiplying the number of all possible combination of y successes in n observations by the probability of y successes in each of the combination as:

$$P\left(y \middle| n, {}^{\lambda}/_{n}\right) = \frac{n!}{y!(n-y)!} ({}^{\lambda}/_{n})^{y} (1 - {}^{\lambda}/_{n})^{n-y} \qquad \text{Equation 16}$$

When n is huge, and p is relatively small. With this assumption, the Equation 16 could then be converted to the form (Jahanshahi et al., 2009):

$$P(y, \lambda) = \frac{\lambda^{y} e^{-\lambda}}{y!} \qquad \text{Equation 17}$$

Take the log gives

$$lnP(y, \lambda) = -\lambda + yln\lambda - \ln(y!) \qquad \text{Equation 18}$$

Poisson regression models are estimated by specifying the Poisson parameter λ_i as a function of explanatory variables X. The most common relationship between explanatory variables and the Poisson parameters is the log-linear model:

$$E[y_i] = \lambda_i = EXP(\beta X_i) \; or \; LN(\lambda_I) = \; \beta X_i \qquad \text{Equation 19}$$

where i is an observation, y_i is the observed number of events, in this research, it is the number of visits to an outdoor recreational destination; X_i is a vector of explanatory variables (e.g., travel time, land uses, land covers, demographic characteristics) and λ_i denotes the predicted number of events per period. There is a basic assumption of Poisson distribution that all dependent variables have the same value for their mean and variance. Cameron and Trivedi (1990) revealed that the distribution of an individual's trip rates for the majority of travel purposes is not in line with the key Poisson regression assumption. The negative binominal regression has been developed to overcome this limitation. The form of the negative binominal regression is written as:

$$E[y_i] = \lambda_i = EXP(\beta X_i + \varepsilon_i) \qquad \text{Equation 20}$$

Where $EXP(\varepsilon_i)$ is a gamma distributed error term with a mean equal to one and variance α^2. The addition of this error term allows the variance to differ from the mean. α is called the over dispersion parameter. Most statistical software nowadays is able to calibrate the NBR model. In this research, the NBR model is calibrated using the R Studio.

The main problem for the NBR has been challenged for many decades by transport modellers as seen in the review in section 2.2: First, this method relies on zoning individuals geographically. Therefore, the question is how does this model count zonal variation? Secondly, and more importantly, what if the correlation between a number of trips and explanatory variables is not linear? In the next section, one of the most recent studies of outdoor recreation using NBR method is reviewed.

2.7 Application of the negative binomial model in the UK National Ecosystem Assessment (NEA) outdoor recreation model

The Poisson and Negative Binomial Regression is a prevalent method in the environmental economics field. For example, Jones et al. (2010) has estimated the number of informal recreational visits to woodland area. Martinez-Espineira et al. (2008) looked at trips to the Gros Morne National Park in Canada. Shreshta et al. (2007) studied nature-based recreation in public areas of the Apalachicola River region in the United States, and Bowker et al. (2007) estimated the economic value of recreational trails in the Virginia Creeper Rail trail. All of these studies were based on on-site observation and applied either the Poisson distribution or the NBR. The only general outdoor recreation research was conducted by Sen et al. (2011, 2014) from the Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia. This study is part of UK NEA, the first analysis of the UK's natural environment in terms of the benefits it provides to society and continuing economic prosperity (Bateman et al., 2014).

The assessment of outdoor recreational activities (Sen et al., 2014) necessitated the development of a methodological framework using off-site household survey data called the Monitor of Engagement with the Natural Environment (MENE). The MENE survey is a questionnaire-based survey conducted by Natural England since 2009. The survey is about how and why people engage with England's natural environment, collects information about the ways that people engage with the natural environment such as visiting the countryside, enjoying greenspaces in towns and cities, watching wildlife and volunteering to help protect the natural environment. Fieldwork started in March 2009 with around 800 respondents interviewed every week across England using an in-home interview format. Every year, at least 45,000 interviews are undertaken[1]. This is the only and most comprehensive survey regarding outdoor recreational trips. This dataset also included primary empirical evidence to build up the new model in this research.

[1] https://www.gov.uk/government/collections/monitor-of-engagement-with-the-natural-environment-survey-purpose-and-results

The modelling method was applied in the UK NEA's study for the negative binomial model. The model was used to predict the number of visits made from each outset location to any given recreational site. The number of visits is assumed to depend on several explantory variables, including land covers of the destinations and alternatives. The alternative was represented using a 10km buffer area around the origin, travel time, demographic information such as the percentages of retired people, the proportions of non-white ethinicity, total population and median of income. Random intercepts are used to catch unobserved correlations; for example, people from the same place may or may not have emotion attached to certain greenspaces. The function was written as:

$$ln(y_{ij}) = y_{00} + y_{01}W_j + y_{10}X_{ij} + u_{0j} + r_{ij} \qquad \text{Equation 21}$$

Where i denotes outset areas (i.e., where the trips start from, defined per the Census as lower super output areas or LSOA in England), j indicates destination sites (represented by the grid cells) and y_{ij} is the predicted number of trips. These trips were from the LSOA *i* to a site *j*. y_{00} is the constant to be estimated. The explanatory variables consist of W_j (which includes variables that describe site characteristics) and X_{ij} (which include variables that describe the outset area characteristics, and travel time). The random part of the model consists of u_{0j} the site-specific random intercept term and hence captures the unobserved heterogeneity between different sites) and r_{ij} (the usual error term). The random effects $u_{0j}$are assumed to be normally distributed with mean zero and variance $\sigma^2\mathrm{u}$.

Table 2.3 *Parameters for Trip-Generation Function (Sen et al., 2014).*

	Coefficients	*t* stat
One-way trip travel time from outset to site		
Travel time (in minutes)	−0.180***	(−159.0)
Land use variables measured at destination		
Log (%Coast at site)	0.158***	(5.950)
Log(%Freshwater at site)	0.0716***	(3.859)
Log (%Other marine at site)	0.0693*	(2.392)
Log(%Mountains at site)	0.0421*	(2.394)
Log (%Woodland at site)	0.0414***	(3.905)
Log (%Grasslands at site)	0.00233	(0.207)
Log (%Urban at site)	−0.224***	(−20.86)
Substitute availability variables measured at outset		
Log (%Coast substitute availability)	−0.0215**	(−2.632)
Log(%Freshwater substitute availability)	−0.0637***	(−6.181)
Log (%Other marine substitute availability)	−0.0325***	(−5.261)
Log (%Mountain substitute availability)	−0.000589	(−0.058)
Log (%Woodland substitute availability)	−0.0622**	(−2.900)
Log (%Grasslands substitute availability)	0.0276	(0.968)
Log (%Urban substitute availability)	−0.441***	(−26.82)
Demographic variables measured at outset		
Log(Median Household Income) (in pounds)	0.430***	(13.19)
% Retired	0.00527**	(2.735)
% Non-white ethnicity	−0.00846***	(−8.623)
Total population of outset area (no. of people)	0.000290***	(6.963)
Constant	−4.458***	(−12.78)
$\ln(\sigma_u^2)$	−0.873***	(−22.29)
σ_u	0.646***	(52.606)
Observations	4,034,290	

The use of this forecasting model is a planning tool for examining the consequences of implementing alternative polices. This model was first applied to test six UK-NEA future scenarios for the UK by 2060 (Haines-Young et al., 2011). Table *2.4* summarises the various changes in each scenario (Bateman et al., 2013). The changes were fed through the UK NEA's model to yield corresponding estimates of visit numbers. The estimated number of trips were further used as an input to calculate corresponding recreational values. Table *2.5* sums these estimates at the national level and calculates the per capita equivalents.

Table 2.4 *Summary of Trends in the UK-NEA Scenarios (Source: Haines-Young et al., 2011)*

Variable	Baseline (2010)	Scenario[a]					
		WM	NW	GF	GPL	LS	NS
% urban	6.7	↑↑	≈	↑	≈	≈	≈
% heathlands	13.8	↓↓	↑↑	↑	↑	≈	↓↓
% grasslands	15.9	↓↓	↑↑	↑	↑↑	↑↑	↓↓
% conifer	5.3	≈	↑↑	↓	↓	↓	↑↑
% broadleaf	6.3	≈	↑↑	↑	↑↑	≈	↑
% farmland	43.5	≈	↓↓	↓	↓	↓	≈
% other	8.3	≈	≈	≈	≈	≈	≈
Δ population	–	↑↑	↑	↑↑	≈	≈	↑↑
Δ real income	–	↑↑	↑↑	↑	↑↑	≈	≈

Variables starting "%" refer to percentages of the total area of Great Britain
↑↑ = proportionally large increase from baseline
↑ = proportionally small increase from baseline
≈ = no substantial change from baseline
↓ = proportionally small decrease from baseline
↓↓ = proportionally large decrease from baseline
[a]Scenario names: WM, World Markets; NW, Nature at Work; GF, Go with the Flow; GPL, Green and Pleasant Land; LS, Local Stewardship; NS, National Security. Δ = change

Table 2.5 *Total (million£) and Per Capita (£) Value of Predicted Annual Visits in the Baseline Period and Changes in Total and Per Capita Value of Predicted Annual Visit under the Various Scenarios (Sen et al., 2014)*

Region	Total baseline (million £)	Change GF (million £)	Change GPL (million £)	Change LS (million £)	Change NS (million £)	Change NW (million £)	Change WM (million £)
England	4,546	2,142	3,917	3,347	2,466	9,056	−1,024
Scotland	1,682	632	1,453	583	688	3,280	−856
Wales	504	179	340	289	291	1,014	−285
GB	6,732	2,953	5,710	4,219	3,445	13,350	−2,165
GB population (millions)	55.4	72.4	62.8	62.0	67.5	65.6	74.5
GB per capita values (£ p.a.)	121	13	77	55	30	185	−60

This model has also been applied at the regional level to optimise the location of new green spaces through estimating the number of trips. This is done via a case study in Northampton, where the analysis identified a particularly suitable area for the recreational site on the north edge of town. The model is also used to estimate the number of visits generated by the creation of the new green space and the expected reduction in visits to surrounding sites (Figure 2.2).

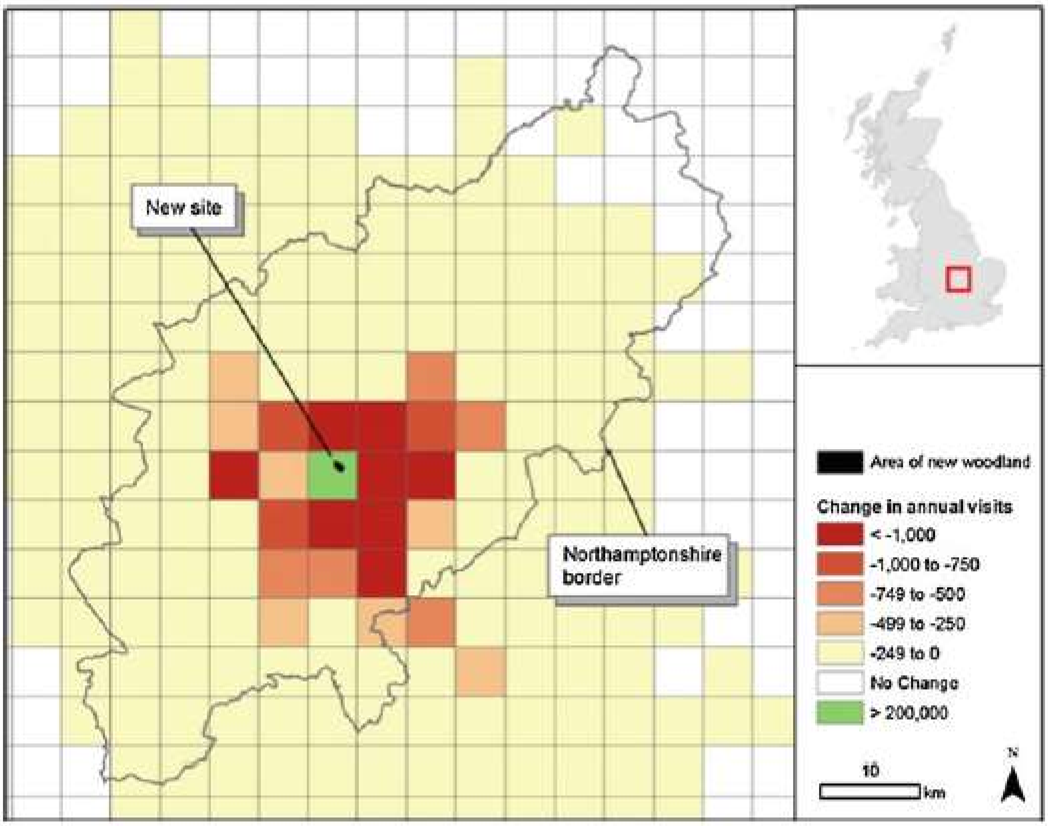


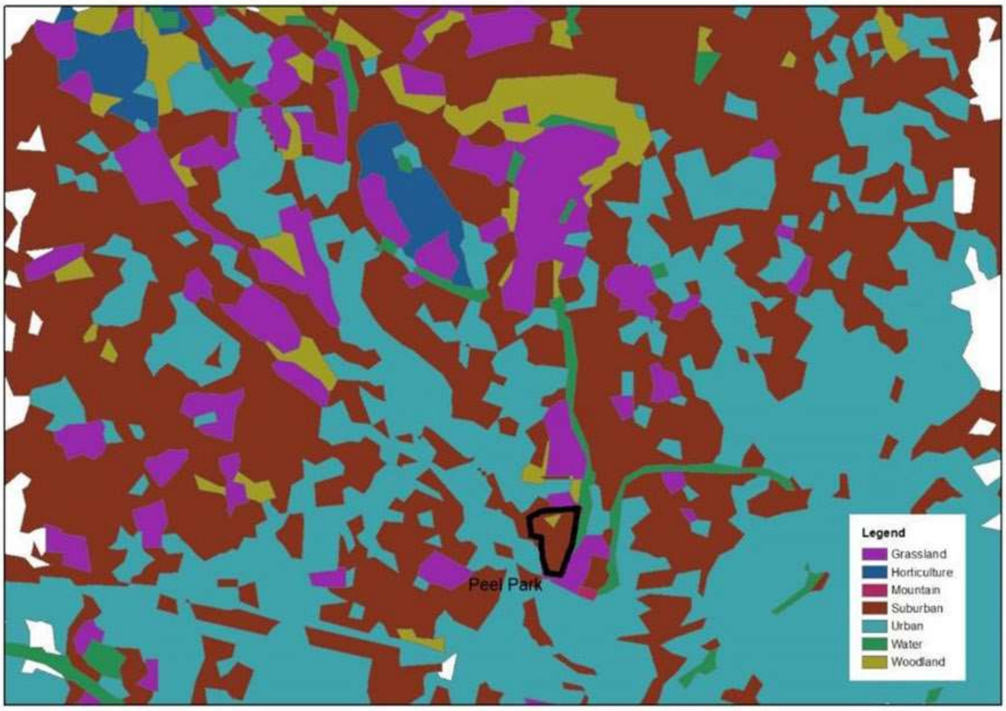
Figure 2.2. Spatial distribution of the estimated change in annual visits to new woodland (Sen et al., 2014).

On the basis of the above reviews, it is not difficult to see that the strengths of the UK NEA's NBR model. Firstly, unlike previous studies which focused on a single site or habitat, this framework can be applied to estimate recreational demand and values for any spatial unit and habitat mix, thanks to the annual in-house survey data from the MENE survey, which has only become available since 2010. Secondly, this model has incorporated environmental characteristics, which are rarely considered by travel forecasting modelling. Thirdly, the applications of the UK NEA's model have revealed that forecasting outdoor recreational trips through environmental characteristics can provide planners with empirical evidence of how people are using green spaces. This kind of model did not exist before, but it is valuable because it can assist decision makers by estimating the changes in value arising from different scenarios at the national level. It is also able to optimise the location of the proposed green space at the local planning level by forecasting the number of visits to the new site.

However, some weaknesses of this method are evident as well. First, as reviewed in the development of transportation modelling, forecasting travel demanding should be a part of the studies of human choice behaviour. It is not theoretically consistent to study choice behaviour purely based on a statistical method such as the NBR model. Studies relying on the statistical method are usually location dependent and are difficult to transfer to different places. Also, the NBR is a zonal model, which means it can only be operated at a zonal level. It is more likely to suffer biases caused by variations among the individuals within each zone.

Secondly, the environmental characteristics of sites are defined by linking their one km square grid cell locations to habitat proportions derived from the 25m resolution UK-wide Land Cover Map 2000 data. The land cover map is produced by the Centre for Ecology and Hydrology[2]. This is a digital map of Great Britain derived from satellite imagery since 1990. The land surface is identified as a collection of discrete irregular objects such as forests, lakes, urban areas and fields using object-based image analysis (OBIA) techniques. Land Cover Map 2000 was derived from image segments and was assigned land cover values according to the pixel distributions within. These classifications were then refined using contextual and ancillary information. Since last time the author has checked, Land Cover Map 2007 is the latest version, built upon the successes of Land Cover Map 2000 and employing similar but enhanced classification techniques. Therefore, the apparent weakness of these data is out of date. Moreover, when zoomed to the neighbourhood level, land shapes, particularly of the open space in the cities, are only partially recognised, and the detailed land cover types are apparently mistaken even in the 2007 version. (See the comparison of what was identified in Peel Park in the LCM2007 with the Google Map image of the same place in Figure 2.3.) In conclusion, the Land Cover Map might be useful when the investigation is at the national or regional level, or in places such as the countryside where changes have been insignificant over the last decade, but it is not helpful when investigating the environmental characteristics of the individual outdoor recreational site at the local level.

[2] www.ceh.ac.uk/landcovermap2007.html



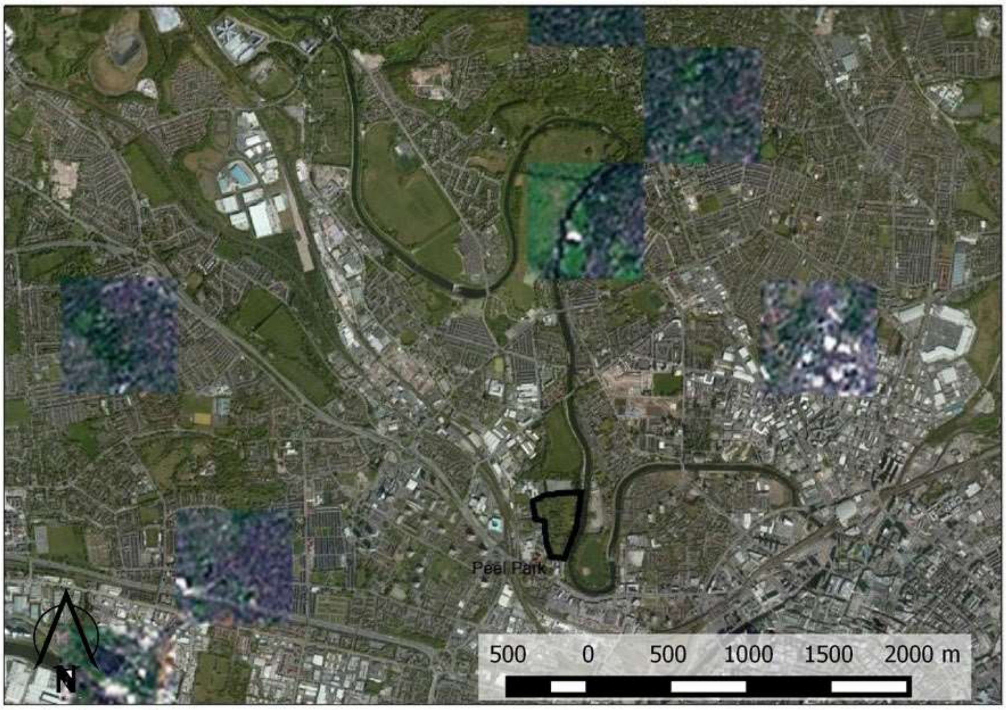


Figure 2.3. Comparison of what was identified of Peel Park in the LCM2007 with the google map image of the same place.

Thirdly, the UK NEA's model includes only land-cover data but has not mentioned anything regarding land use. Land-use data is more relevant than land-cover data when planning decisions are made. This weakness significantly limits the application of the model.

Travel time in studies carried out by Sen et al. were calculated using the Ordnance Survey Meridian road network, and average road speeds from Jones et al. (2010). The study by Jones et al. (2010) assigned a speed to a type of road (e.g. Motorway, A-road, B-road and minor road), and it also discriminated the differences between urban and rural contexts. The road network is converted into a regular grid of 100 x 100 metre cells with each cell containing a value corresponding to travel-time-per-unit distance. The resultant travel time map is used to calculate the minimum travel time between any outset location and any destination site (Sen et al., 2011). The noticeable problem with this method is that the assumption disregards travel mode and road congestions, which are considered essential when estimating the cost of travel.

As a result, although the UK NEA's model gives a fair estimation on trip accounts at the national and regional scale, it faces various challenges when estimating trips to an individual outdoor recreational site for the reasons discussed above. Consequently, it is not expected to be used in making estimations of visits to a single destination, and this shortcoming was confirmed by personal communication with Dr. Antara Sen: 'please note that our recreation model was designed specifically to predict visits on a national scale. We did face a number of problems when applying it to a local scale study'.

2.8 Summary

Random Utility Maximising (RUM)-based Discrete Choice Models (DCMs) have been well-developed during the last 50 years and applied extensively for forecasting peak-time journeys to work and business trips in the transportation field. However, the outdoor recreational travel demand remains poorly understood in transportation modelling because, first, outdoor recreational trips usually take place in the off-peak time outside commuting rush hours and are thus unlikely to be a leading cause of traffic congestion, which is the primary focus for transportation modellers. Secondly, before the MENE survey, it lacked observation data for people to study general outdoor recreational activities. Previous studies on outdoor recreational trips are mainly based on on-site observation with a focus on single site or habitat, and the results are difficult to be transferred to any other place. Thirdly, the MENE survey so far has only been used by the UK NEA's study, through applications of the NBR (Negative Binomial Regression) model. Although the UK NEA's research has proven the value of this kind of model for city planners and designers by testing different planning scenarios, the model has not been designed to predict the effects of interventions at the individual site level.

In conclusion, there is a distinct gap in our knowledge and analysis of outdoor recreational travel. Building a new travel demanding model for outdoor recreational trips will be necessary to fill this gap, and this will be achieved by answering the following four research questions: First, how to build the new model for outdoor recreational travel? Secondly, is the estimation accurate enough? Thirdly, to what extent can the new model be transferred to destinations outside the case study area? And finally, how can city planners and designers use this new method?

Chapter 3 Overview of the Analytical Process

3.1 Overview

The motivation for this research is to develop a method that can assist decision makers with robust evidence on what attracts people to take outdoor recreational trips, so that their development decisions can be made without harming the benefits of engaging with the natural environment. The way to achieve this is through developing a new travel demand model as reviewed in Chapter 2. Therefore, the aim of this research is to build a new travel demand model for outdoor recreational trips. The questions raised at the end of the last chapter will be answered through the conventional transport modelling process, which includes data collection and preliminary analysis, model calibration, model validation and scenario tests (Figure 3.1). The first two steps are described in Chapter 4. Model validation and scenarios tests will be presented in Chapter 5 and Chapter 6 respectively. This chapter is an overview of how Chapters 4 to 6 will fit together to answer the research questions.

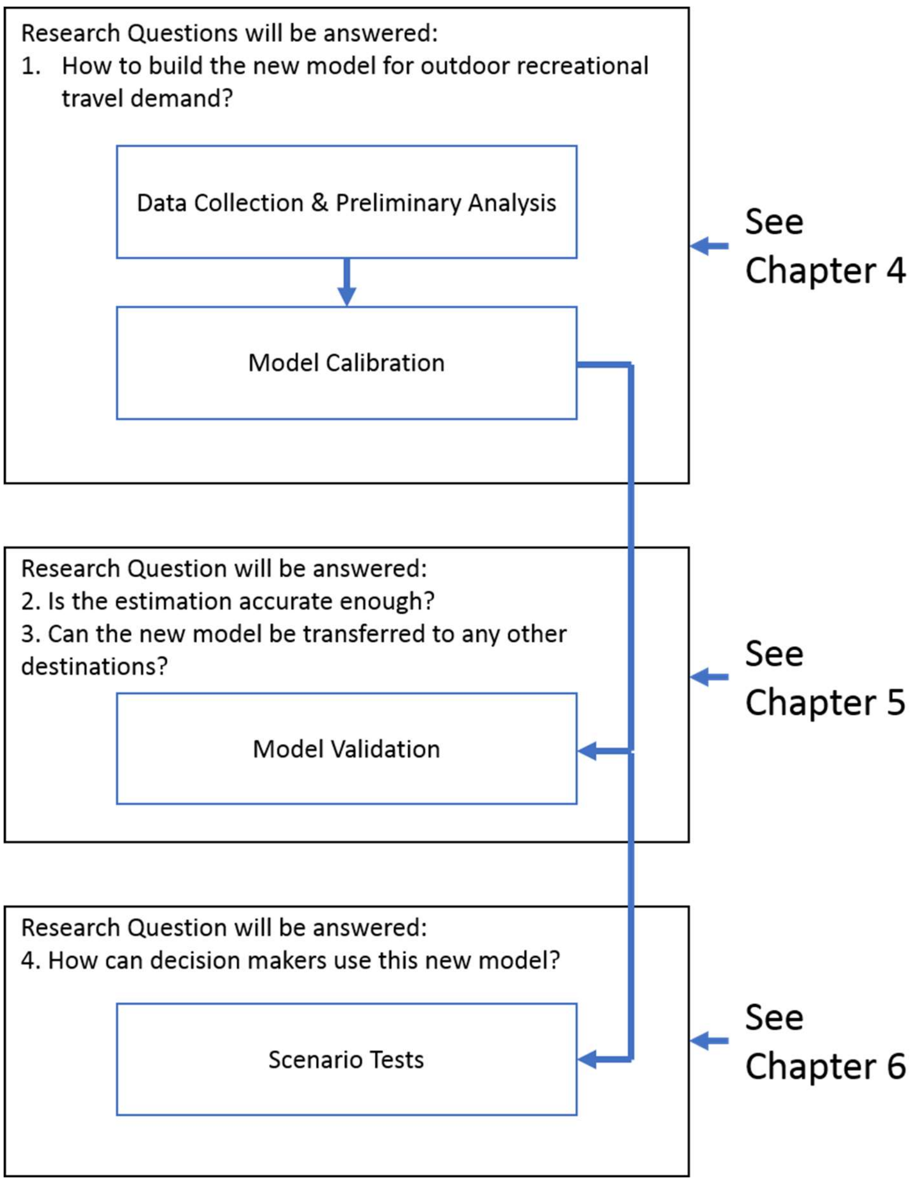


Figure 3.1. Modelling process.

3.2 Data collection and preliminary analysis, model calibration

As depicted in Figure 3.1, the question to be answered in Chapter 4 is how to build a new discrete choice model for outdoor recreational trips. This is done through a case study on a selected area, which covers 14 local authorities. The model-building process follows standard modelling procedure as shown in Figure 3.1. Firstly, there is data collection and preliminary analysis. Various data were collected for this area to build the new model. The variables includs demographic data from the ONS; origin and destination information from the MENE survey; travel time captured from Google directions API; and environmental characteristics derived from combining data from OpenStreetMap, the Generalised Land Use Database (GLUD); and the MENE survey.

The new model follows the conventional transportation model structure as reviewed in Section 2.2. Firstly, the trip generation function calculates the total number of trips from each origin zone. Then, it turns to the mode choice and distribution functions, which are in the form of discrete choice models. The model training process was divided into two stages. Firstly, define the model structure. The data were firstly run through a multinomial logit model and then tested in the nested logit models. When tested in the nested logit model, two types of structure were applied: the first one is based on the assumption that the individual chose the travel mode before the destination. Secondly, the other way around. The better structure is decided on the basis of the restriction of the nested logit model that parameters associated with different levels should not increase under progression towards the top of the tree structure (see Section 2.3.2). The tests using nested logit model also investigated whether the IIA (see Section 2.3.1) assumption is valid held in the multinomial logit test. In the last step of the first model training stage, the data were also tested through the mixed logit model to investigate the variations between individuals. Although the final model form will not be in the mixed logit form, the coeefficents of variables are random numbers following the normal distributions, and they are difficult to explain for planning purposes. But it is useful to be aware of where the variations happened, and it might require being studied separately in the future.

The second stage of building up the new model is finalising the explanatory variables. This is done by testing different combinations of variables through the same model structure. The structure was decided according to the above steps. The ultimate variable combination is judged by the best statistical model results. Three experiments were conducted: firstly, replacing the GLUD variables with a single variable indicates the area of green space; secondly, implementing distance limits when considering the alternatives of destinations; and, finally, dividing the data into groups by activity. After these two stages, the new model was finalised.

3.3 Model Validation

Chapter 5 concerns model validation, which answers the next two questions: is the new model making an accurate estimation, and can this model be transferred to any destination in England? Both questions are answered by comparing the model estimations with the records of visits from independent sources. The first research question is answered by two tests: one on Wigg Island and another on Wigan Flashes. These are two nature reserves inside the model calibration area. Management teams on both sites have recorded the usage of each reserve in different ways. The new model was run to calculate the number of trips for each nature reserve. Then the predictions were compared with the observations. Once it had been proven that the model can make comparable results for these two destinations, the next estimation was made for the ten English National Parks. The estimations from the model were compared with the number from the National Country Park report (2013). This is to answer the question: can this model be transferred to any destination (irrespective of size) in England?

3.4 Scenarios tests

Through answering above two questions, the new model can be trusted and transferred to any outdoor recreational destination in England. The last question of this research is: how can city planners and designers use this new method? Inspired by the UK NEA model, the new model is applied to test three different scenarios for the Upper Mersey Estuary (UME) area: Business As Usual (BAU), Development Boom (DB) and Nature is the Key (NK). These scenarios have been developed by Dr Andrea Drewitt of the Ecosystems and Environment Research Centre at the University of Salford. Different land use strategies are proposed on the basis of the primary development focus as their names suggest (See Chapter 6). Model results are presented in the ways of the changes of visits to three recreational sites inside the UME area: two of them are existing nature reserves (Wigg Island and Moore Nature Reserve), and another one is a proposed new country park (Arpley Country Park).

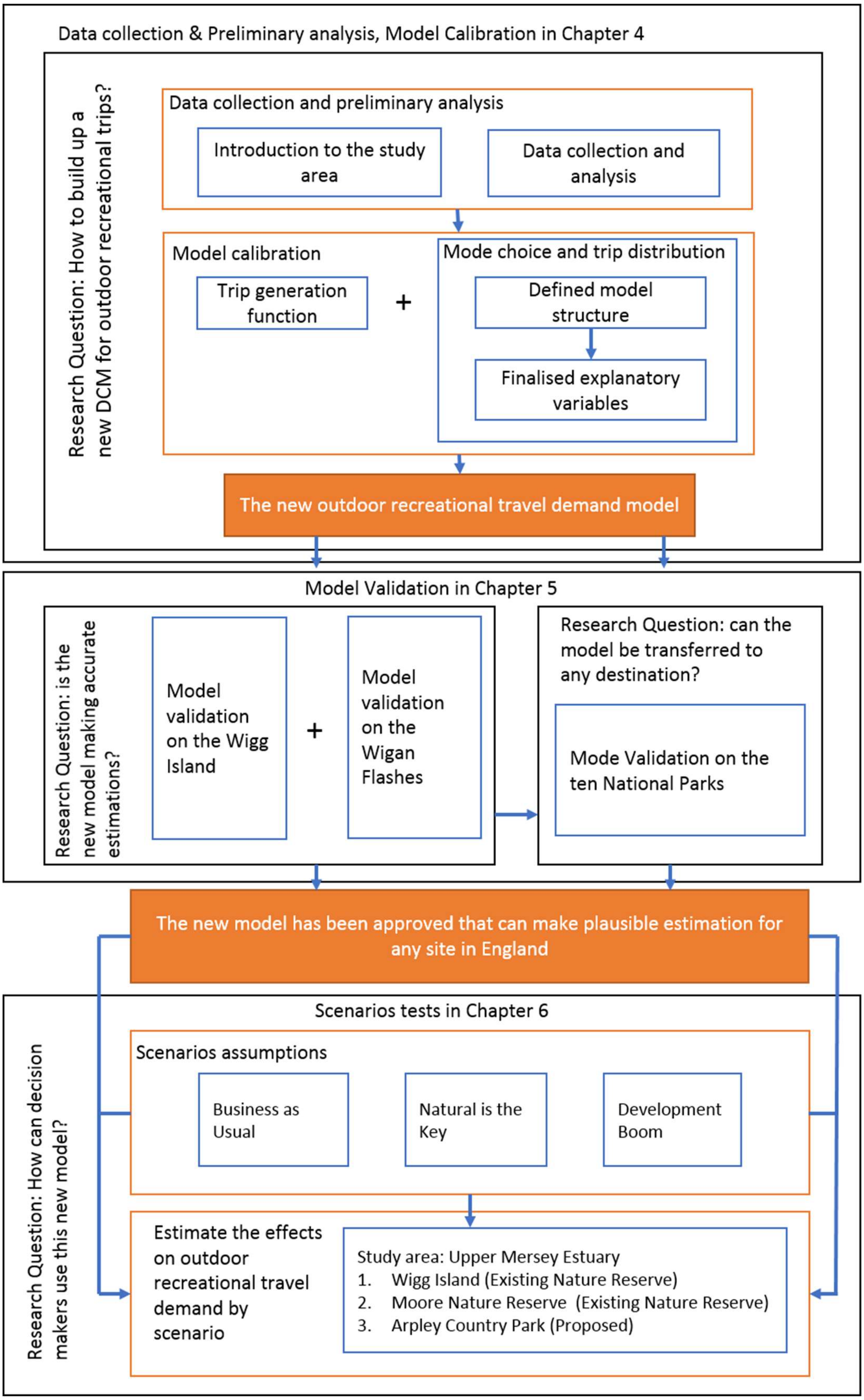


Figure 3.2. Diagram shows the structure of Chapters 4 to 6.

3.5 Summary

Research Questions:	Steps	Case studies:
How to build the new model for outdoor recreational travel demand?	**Data collection & preliminary analysis** **Model Calibration (including three phases)** **Trip generation** looks at how many outdoor recreational trips people will make per year from their home. Result: number of trips from each origin Question: how would these trips be distributed to different transport modes and destinations? **Mode choice**: shows which mode people use to travel between their origins and destinations Result: Number of trips by mode from each origin. Question: How would these trips be distributed to each of destinations? **Trip distribution**: looking at where each trip goes. Result: number of trips to each destination. Question: is the estimation accurate enough?	14 selected districts in the North-West region
Is the estimation accurate enough?	**Validate the new model**: check whether the new model makes a plausible estimation of the number of trips to individual site, Result: the new model produces a plausible estimation. Question: can this model be transferred to any other site in England?	Wigg Island and Wigan Flashes
Can the new model be transferred to any destination?	**More validations** Result: the new model can be transferred to any other site in England. Question: how to use this model?	National Parks
How can planners use this new method?	**Application of the new model**: scenarios tests	Wigg Island, Moore nature reserve and Arpley country park

Table 3.1. *Main Research Questions, Empirical Study Processes, and the Case Study Areas*

This chapter has depicted the main outline of Chapter 4 and Chapter 5, which contain the analytic process for this research. These studies are used to answer the four research questions raised at the end of Chapter 2 (Table *3.1*). From the next Chapter, the details of the analytic process will be reported step by step.

Chapter 4 Model calibration

4.1 Overview

The aim of this chapter is to answer the first research question: how to build the new model for outdoor recreational travel? This is done through training a new outdoor recreational travel demand mode on a selected case study area. The case study area for model calibration is introduced in Section 4.2. The variables used in this study are informed by previous research, and these studies are reviewed in Section 4.3. Data collection methods and statistical summaries of the variables are presented in Section 4.4. The process of building the new travel demand model is described step by step starting from section 4.5. The best form of trip-generation function is presented in Section 4.5. In the first part of Section 4.6, the structure of modal choice and trip-distribution functions are explored through a series of tests on the different forms of Discrete Choice Models (DCMs). And, in the second part of 4.6, different sets of exploratory variables are examined in order to work out the best model forms for trip-distribution functions. The final model is presented in Section 4.7, and finally, the chapter is summarised in Section 4.8. The structure of this chapter can also be seen in Figure 4.1.

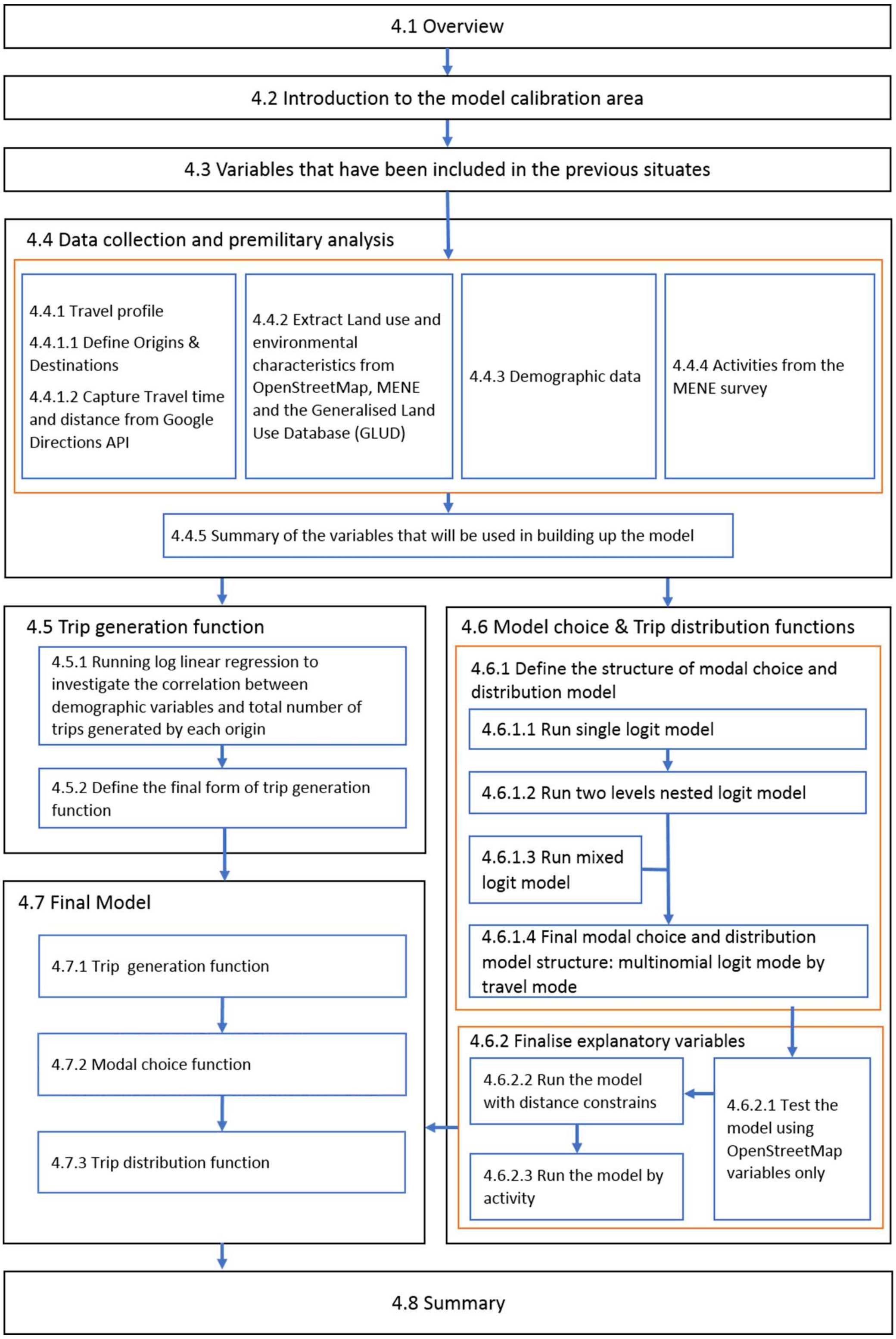


Figure 4.1. Chapter 4 structure.

4.2 Introduction to the model calibration case study area

The study area used for model calibration involves two ceremonial counties in the North-West Region. It covers all of the Cheshire county and six of the ten districts in the Greater Manchester area (Table 4.1). Cheshire is a county bordering Merseyside and Greater Manchester to the north, Derbyshire to the east, Staffordshire and Shropshire to the south and Wales to the west. The county covers 2,343 square kilometres, and, according to the 2011 census data, it has a population of around 1 million. The population density is 32 people per square kilometre. It is lower than the North-West average of 42 people per square kilometre. It is mostly rural with some small towns and villages supporting the agricultural and other industries.

Greater Manchester is a combined authority in North-West England. According to 2014 mid-year estimations from the Manchester city council[3], its population has reached 2.8 million on a 1,276 square kilometres area. There is a mix of high-density urban areas, suburbs, semi-rural and rural locations in Greater Manchester, but land use is mostly urban. It has a focused central business district, formed by Manchester city centre and the adjoining parts of Salford and Trafford, but Greater Manchester is also a polycentric county with ten metropolitan districts: Bolton, Burry, Oldham, Rochdale, Stockport, Tameside, Trafford, Wigan, and the cities of Manchester and Salford. The six western boroughs are included in this study; excluded are the four boroughs on the eastern border: Oldham, Rochdale, Stockport, and Tameside.

The boundary of case study area was drawn as shown in Figure 4.2 for two reasons: firstly, this coverage facilitates in-depth studies in green space scenarios (see Chapter 5). Secondly, 3501 interviewees from the MENE data were collected within this boundary. This has given us sufficient samples to carry out further analysis and training the new model.

---

[3]http://www.manchester.gov.uk/downloads/download/4220/corporate_research_and_intelligence_population_publications

Table 4.1 *Upper Tier Local Authorities Included in Research Area*

Upper Tier Local Authorities
Cheshire
Cheshire East
Cheshire West & Chester
Halton
Warrington
Wirral
Greater Manchester
Bolton
Bury
Manchester
Salford
Trafford
Wigan

Contains Ordnance Survey data © Crown copyright and database right 2015

Figure 4.2. Case study area.

4.3 Variables included in previous studies

The first step of building the new model is deciding what the variables are. This is based on a review of previous studies. As previous studies have suggested, a number of variables have been found to play significant roles in outdoor recreational travel demand. First, all studies include a variable to represent the cost of travel, and they all have found that travel is significantly related to outdoor recreation activities.

Travel time by far is the most popular form (e.g., Jones et al., 2010; Sen et al., 2014). Other studies have used distance (e.g., Bestard & Font, 2010; Herriges & Phaneuf, 2010). The problem of studies based on travel distance is that they disregard transport modes. Travel cost is another one of the most popular variables (e.g., Bowker et al., 2007; Francis & Martínez-espiñeira, 2012). Travel cost is usually calculated by multiplying travel time by single unit time cost. Studies have different preferences on the value of single-unit time cost, and they have not reached agreement as to what value should be used (Fezzi et al., 2012; Hagerty & Moeltner, 2005). Therefore, in this research, travel time and travel distance are the only variables tested.

In recent studies the crucial role of environmental characteristics has been highlighted, however, different studies have examined environmental attributes through various forms. Land use and land cover are the most frequently used (e.g., Jones et al., 2010; Paracchini et al., 2014; Sen et al., 2014). Land use is the function of the land. Land cover focuses more on the physical characteristics of the area. Using land cover data needs to include its accuracy, particularly at the small scale such as the neighborhood level. Because the land cover map is derived from the satellite images, accuracy is weak at the local level as noted in Section 4.2.4. In order to avoid this bias, a few studies applied dummy variables to indicate the standard of naturalness (e.g., Joyce & Sutton, 2009; Shreshta et al., 2007).

The third group of variables that have been included in previous outdoor recreation studies is demographic characteristics. For instance, age, income, ethnicity, sex, education level and household size are the most mentioned variables in earlier studies. However, results are not consistent across different studies. For instance, Shreshta (2007) suggested education level was a significant predictor of recreation trips to the Apalachicola River region in Florida. However, Tuffour (2012) found that education attainment is insignificant for the Gros Moren national park in Canada. Following the latest study of general recreational trips by UK NEA (Sen et al., 2013, 2014), the demographic variables tested in this research include percentages of retired people, proportions of the non-white population, the median of income and population.

Last but not least, some variables related to activities are included in many previous studies. For instance, Bowker (2007) used a dummy variable, which equals to one when an individual went to Virginia Creeper Rail Trail for biking, zero for any other activities and found the trail is apparently unattractive to bikers. Herrings and Phaneuf (2010) used a dummy variable indicating ownership of hunting or fishing licenses, and found it significantly increasing the likelihood of taking a trip to any site on the Iowa Wetland. The means of including activities in this research was by grouping the MENE survey data into four groups of activities (see section 4.2.6) and then running the model separately based on each group of data.

The data used in this study are shown in Table *4.2*. In Section 4.2, each will be looked at regarding how each type of data was collected and what they can tell us from an initial analysis.

Table 4.2 *Variables Used in this Study*

Variables
Travel Profile:
Mode
Time
Distance
Environmental Characteristics:
Land use
Land Cover
Demographic:
Population
Percentage of retired population
Percentage of non-white ethnicity
Income
Activities
Walking without a dog
Walking with a dog
Informal sports and Play
Others

4.4 Data collection and preliminary analysis

4.4.1 Cost of Travel

As discussed above, the cost of travel to outdoor recreational sites was investigated in the forms of travel time and travel distance in this research. In this section, the Origin and Destination are firstly defined, and then the focus moves to how the travel time and distance were collected and what the emerging patterns are.

4.4.1.1 Origins and Destinations of trips

The starting point of each trip is the individual's residential neighbourhood. The finest level of information available in England is called the Lower Super Output Area (LSOA). The population weighted centre of each LSOA area represents the origin for people who say their trip started from home. This centroid is different from a geometric centroid because it represents how the population at census time was spatially distributed and grouped within the LSOA[4]. The MENE survey gives us information on whether the trip started from home (Question 9: Did this journey start from your home/someone else's home, work, holiday accommodation somewhere else). If someone said they departed from their home, the population-weighted centroid of the LSOA of their home address was used as the origin. This study only focused on the home-based trip since it is impossible to know the origin for those who began their journey from anywhere else. Destinations of the sampled trips have been documented and geocoded (X, Y coordinates) by the MENE survey team. The scatter plot graph of the destinations is shown in Figure 4.3. This is an image of outdoor recreational destinations spotted by people living in the districts which are included in Table 4.1. Therefore, small and informal green spaces are only identified as destinations where they are close to the study area. Outside the North-West region, only major natural spots in England have emerged on the map, for example, the Lake District, York Moors National Park, and Cornwall.

Since the destinations are recorded in the form of the postcode in the MENE data, one situation happened quite often as shown in Figure 4.4: a single green space has been identified several time at different locations (red points in Figure 4.4), and some destinations located away from the actual green spaces. A unique destination point for each green space (yellow points in Figure 4.4) were identified by combining information from the MENE survey, the OpenStreetMap and web page searches, as described in Section 4.4.2.4.

---

[4] http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/geography/products/census/spatial/centroids/index.html

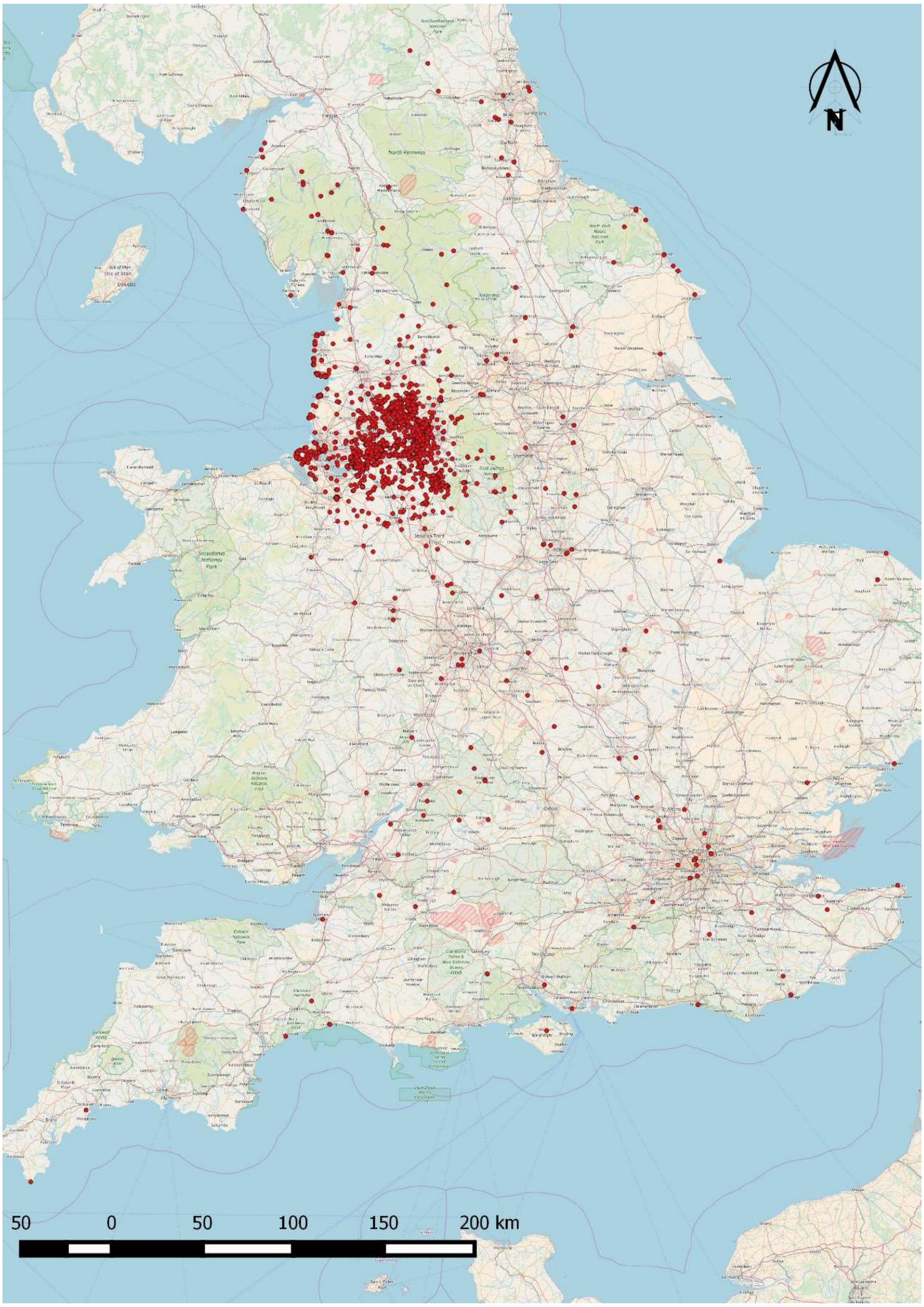


Figure 4.3. Scatter plot of destinations.

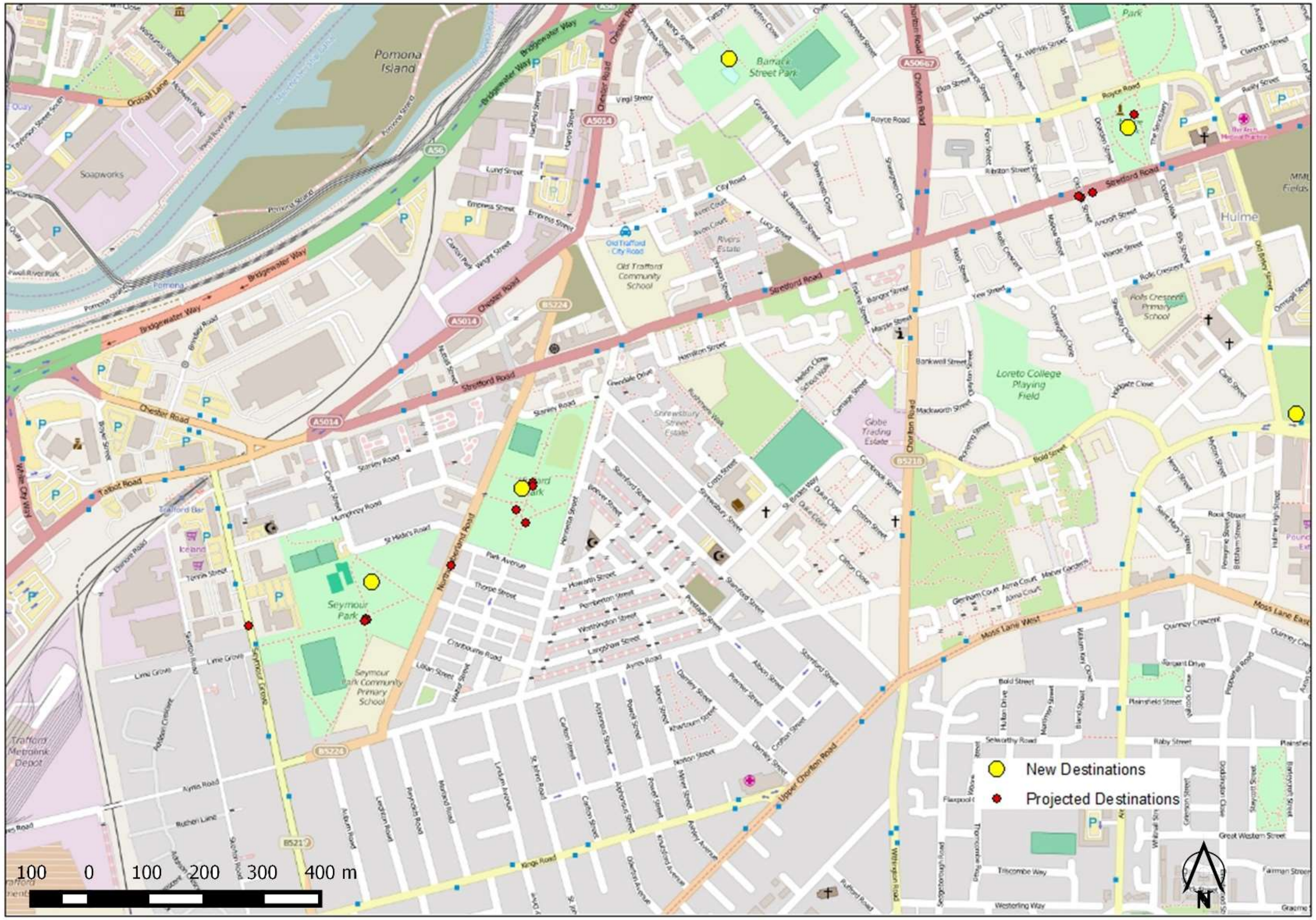


Figure 4.4. Pinpoint destination points (Yellow dots) based on observed destination projected through geocodes.

4.4.1.2 Travel times and distances

The tool for generating travel times and distances is the Google Directions API. The API allows the retrieval of predicted journey information, including the assumed shortest route, and trip duration based on selected transport mode. Four different transport modes can be chosen: driving, walking, cycling, and transit (public transport). Within this model, Google generates real-time traffic flow using crowd-sourced data. Google also receives up-to-date public transport timetable and delay information from TfL and Network Rail[5]. In conclusion, the travel time used in this research is more realistic than in previous studies, wherein the travel time was calculated through GIS tool based on average speed assumptions (for detail reviews see section 2.6).

The observed trips in the MENE survey were recorded in nine different travel options. Those options have been regrouped into the four modes Google API supported as shown in Table *4.3*; travel time was estimated at times at a point in the future calculated by long-term means by day of the week and time of day. In this case, all trips were collected for Saturday, 16 January 2016 as a typical off-peak travel condition[6].

Table 4.3 Regrouping Travel Mode

Google Mode	The MENE Mode
Driving	Car or van, coach trip/ private coach, motorcycle/ scooter, taxi
Walking	On foot/ walking, wheelchair/mobility scooter
Cycling	Bicycle/ mountain bike
Transit	Train (includes Tube/underground) Public bus or coach (scheduled service)

There are very significant differences in the distribution of trips among different groups which are organised by the transport mode. Figure 4.5 depicts a summary of

[5] https://developers.google.com/maps/documentation/directions

[6] These data were collected on December 2015. Saturday 16 January 2016 was chosen because outdoor recreation is more likely to happen during off-peak time. 16 January 2016 is on Saturday and after the Christmas & New Year holiday.

travel times for a single trip to outdoor recreational sites as recorded in the MENE survey. Seventy-five percent of cyclists and walkers spend less than 50 minutes going to outdoor green spaces. This is slightly shorter than individuals travelled by car. Three-quarters of people drove less than 80 minutes for outdoor recreation purposes. Medians for these three modes are 9.5, 13.4 and 18.7 minutes respectively. People who chose to use public transportation have made an even more different pattern: the range of time they spend on the journey is broad, from twenty minutes to three-and-a-half hours. The median for transit is 49.3 minutes. Three of these four modes contain significant outliers except cyclists. The longest trip by public transport took more than 10 hours. There are travel time biases caused by either mistaking where people departed, or the proxy of origin (population-weighted centre) is too far from where people actually live. The way to validate these travel times is comparing the shortest route suggested by the Google Direction API with the data from the MENE survey (Question 8: Approximately how far, in miles, did you travel to reach this place?) In situations where it is difficult to reconcile the two sources, the observations are left out of our dataset.

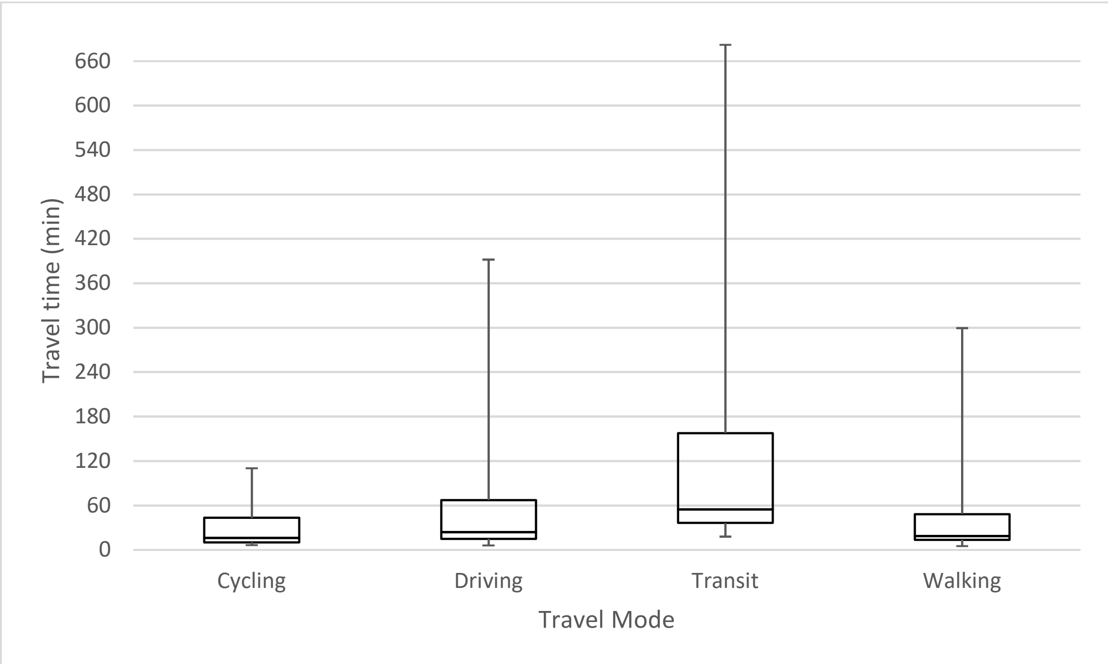


Figure 4.5. Travel time by mode in minutes.

Figure 4.6 illustrates the distribution of the journey distance for outdoor recreations in England, and

Figure *4.7* shows the pattern for case study area. The study area gives a similar travel distance pattern as it is at the national level. Majority people (above 80%) will not go more than ten miles (16 km) for outdoor recreation purposes. Trips these are less than one mile (1.6 km) represent the largest part of both charts (40%).

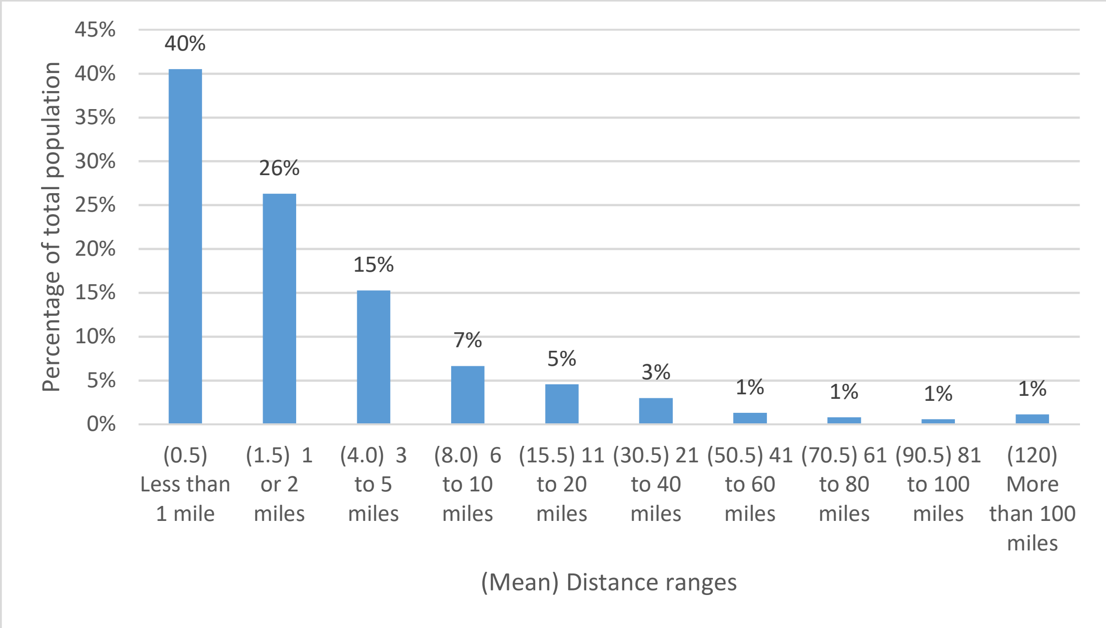


Figure 4.6. Travel distance distribution in England.

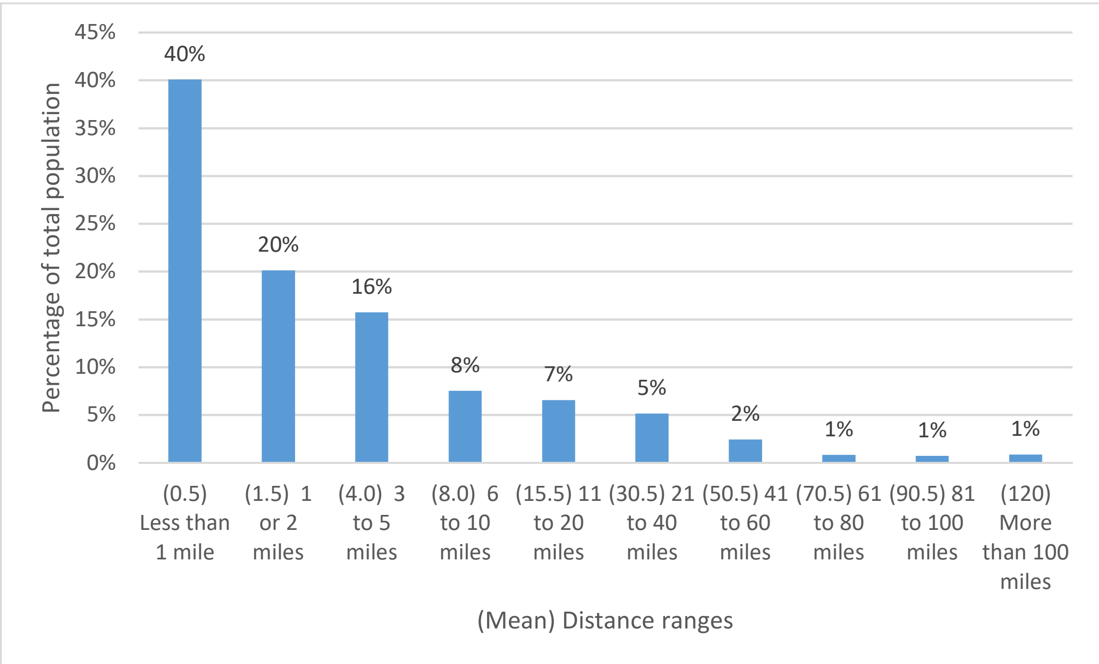


Figure 4.7. Travel distance distribution in study area.

These data have been further divided by different transport modes for analysis. People travelled by different mode, giving us a significantly different pattern in terms of how far/long they would travel for a recreational purpose. For cycling trips (Figure 4.8), up to 65% people traveled less than five miles (eight kilometres), another 20% to 25% people went more than six miles (9.6 kilometres), but less than 40 miles (64 kilometres); very few people moved beyond this distance by cycling. As to driving trips (Figure 4.9), only 7% of individuals chose to drive within one mile (1.6 kilometres). Moreover, 45% of people hit between one to five miles for outdoor recreational trips. Another 30% of people would drive up to 40 miles, and about 20% would drive further than 40 miles, with 5% of individuals going further than 100 miles (160 kilometres). Transit trips give similar patterns to driving tours, as shown in Figure 4.10. The only difference is fewer people used public transport between one to two miles (3.2 kilometres). In addition, 90% of walking trips are less than two miles (Figure 4.11), of which 60% are shorter than one mile. Only fewer than 5% of individuals would walk between three to five miles for the recreational purpose. In conclusion, it is necessary to deal with transport modes separately, which has never been done in previous studies.

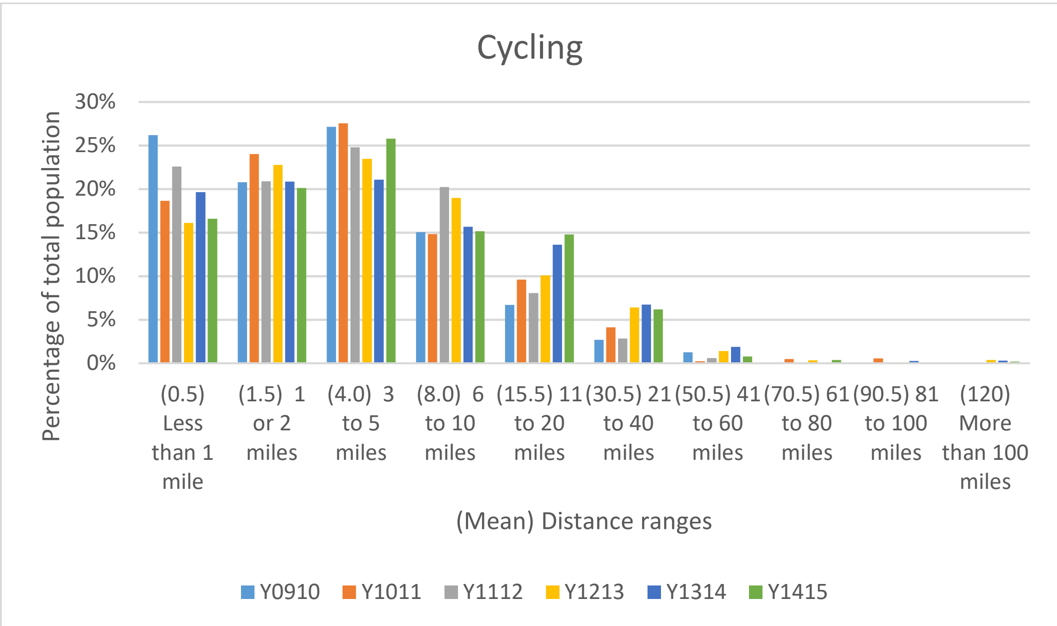


Figure 4.8. Travel distance by cycling.

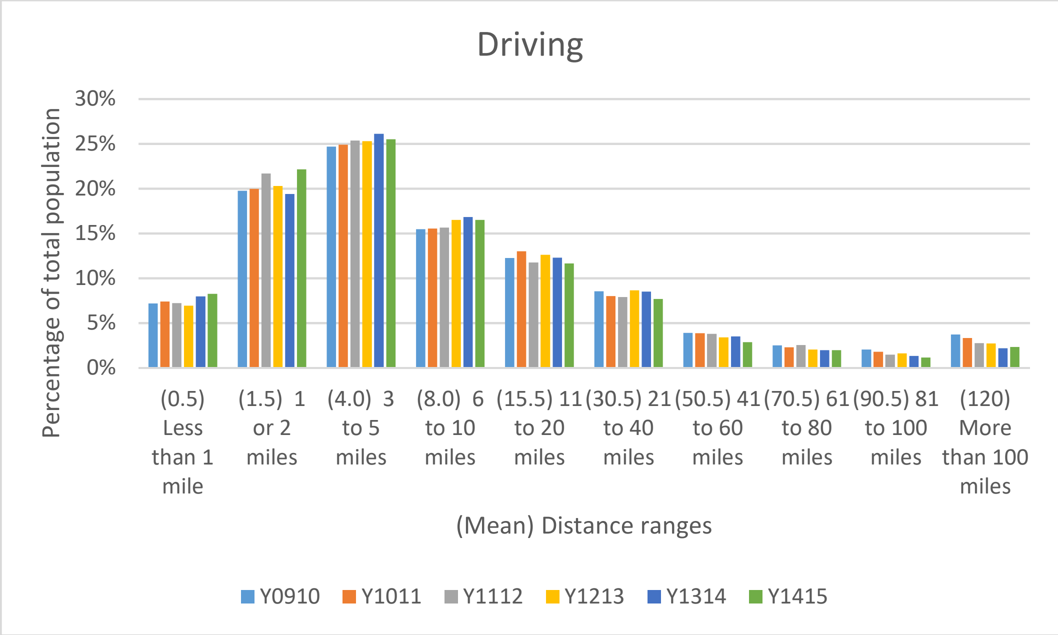


Figure 4.9. Travel distance by driving.

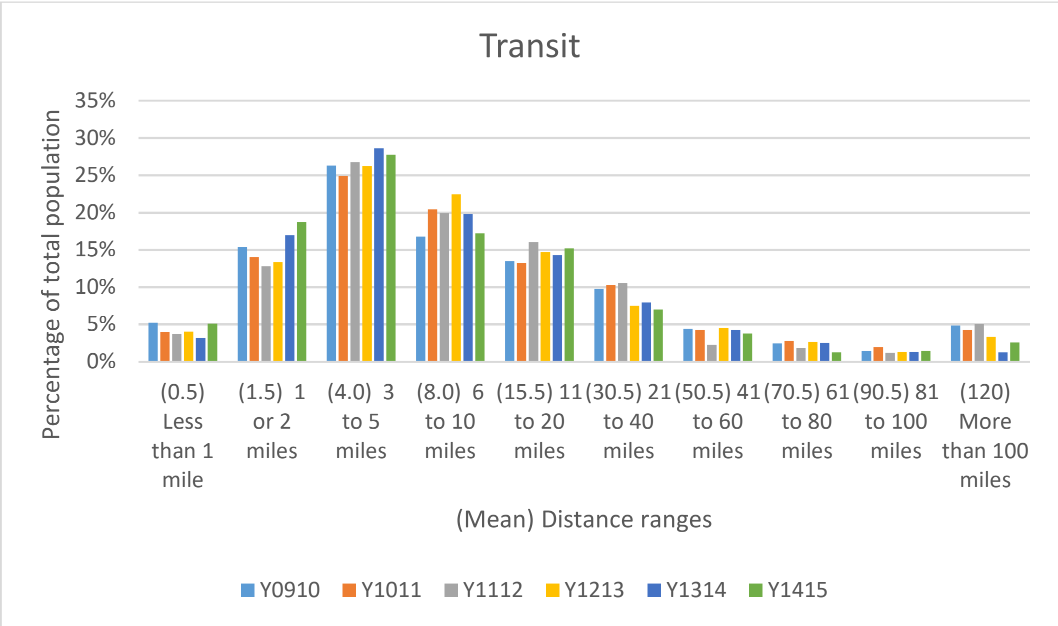


Figure 4.10. Travel distance by transit.

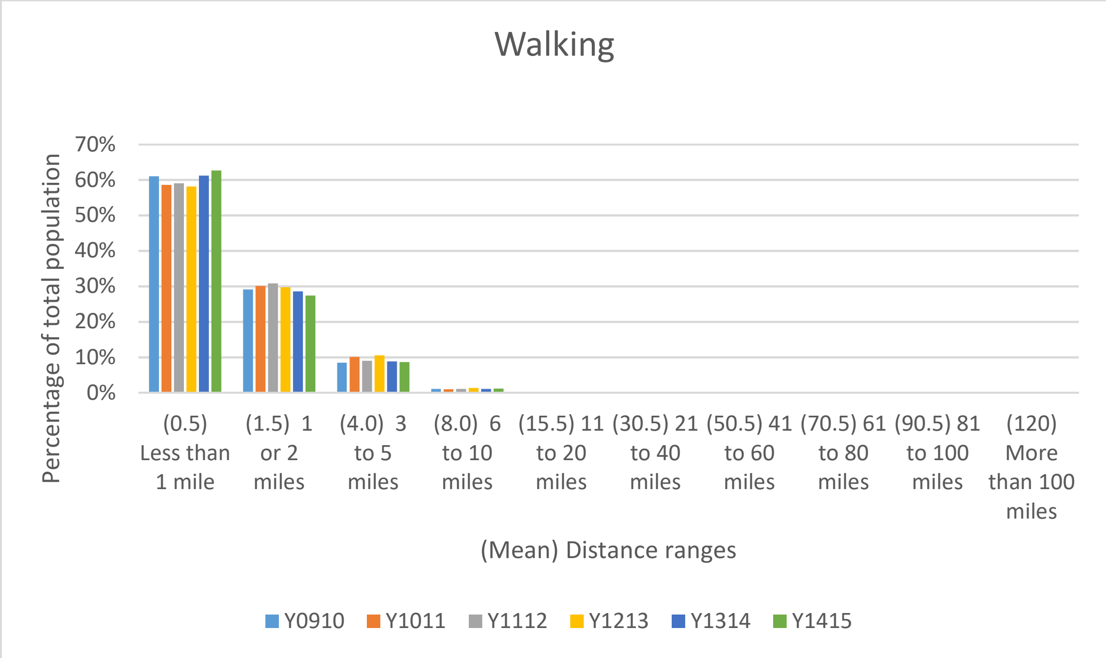


Figure 4.11. Travel distance by walking.

4.4.2 Land uses and environmental attributes

As reviewed in Section 4.2.2, one popular method used to investigate environment characteristics is through land cover data. However, it is remarkably difficult to get this information with precision.

4.4.2.1 Land cover map

A land cover map including all UK areas is available online, produced by the Centre for Ecology and Hydrology[7]. This is a digital map of Great Britain derived from satellite imagery since 1990. The land surface is identified as a collection of discrete irregular objects such as forests, lakes, urban areas and fields using object-based image analysis (OBIA) techniques. Land Cover Map 2000 was derived from image segments and was assigned land cover values according to the pixel distributions within. These classifications were then refined using contextual and ancillary information. Land Cover Map 2007, the latest version available at the time of this study, was built upon the successes of Land Cover Map 2000 and employs similar but enhanced classification techniques. This map provides complete information on the land cover all over the UK. However, when zooming in at the neighbourhood level,

[7] www.ceh.ac.uk/landcovermap2007.html

land shapes, particularly of the open space in the cities, are only partially recognised, and the detailed land cover types do not always appear to be accurate. (See the comparison of what is identified of Peel Park in the LCM2007 with the google map image Figure 4.12.) In conclusion, these data are useful in situations such as investigations at the national or regional level, or in the countryside where changes are insignificant during decades, but they are not helpful in this research because the focus is on local outdoor recreational destinations.

On the other hand, land use is more directly connected to planning strategies than land cover. Research using land cover will have to transfer land cover to land use, in order to be applied in the urban planning process. Therefore, environmental characteristics in this research are in the form of land use, derived through combing information from the OpenStreetMap (OSM) and the MENE survey.

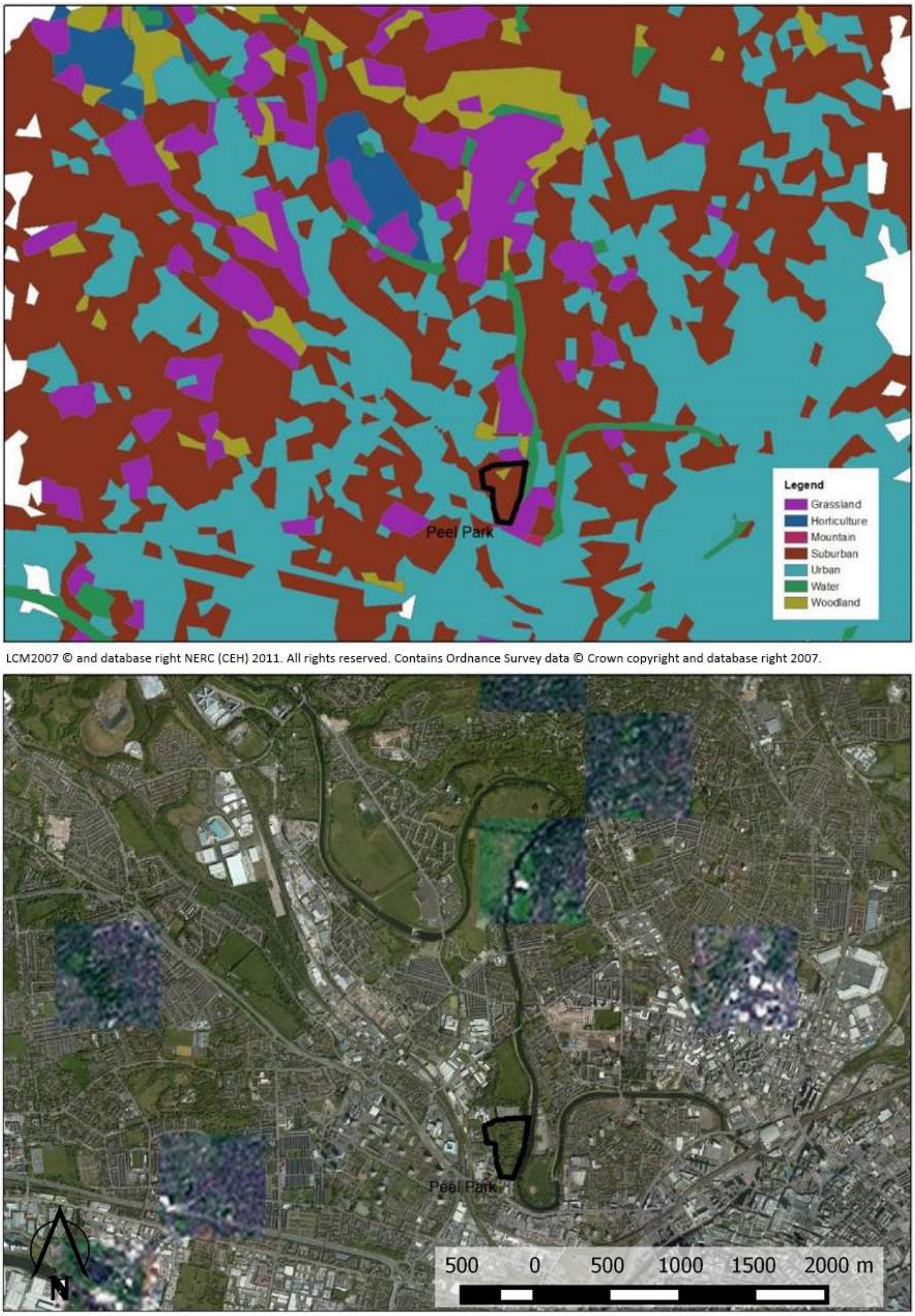


Figure 4.12. Comparing Land Cover Map 2007 with Google Map, Salford.

4.4.2.2 Land type from the OpenStreetMap

OpenStreetMap (OSM) is built by a community of mappers who contribute to and maintain data all over the world. Contributors use aerial imagery, GPS devices and low-tech field maps to create and verify that map data are accurate and up to date[8]. The dataset is available in the GIS shapefile as lines, dots and polygons. Polygons containing the characteristics shown in Table *4.4* are selected. These texts are presented here as what they have been recorded on the OSM. The selected polygons are believed to be the most relevant to outdoor recreational trips.

Table 4.4 *Selected Land Use Types from OSM*

agriculture	caravan_park	flaying field	meadow	public_square
allotments	caravan_site	flowerbed	mountain	recreation_groun
animal_enclosure	carpark	Forrest	nature_reserve	riverbank
animal_husbandry	cemetrey	garden	nursery_horticul	school playgroun
animal_keeping	church	golf	open_space	shrubs
aquaculture	civic_amenity	grass	orchard	sport
artificial_turf	college_court	graveyard	ornamental	suqare
basin	conservation	green_space	park	town_square
beach	courtyard	greenfield	picnic_area	travellers camp
beer_garden	crop_rotation	greenhouse	pitch	vegetable
boatyard	dissused_recreat	harbour	plants	village
brownfield	dock	Hillside	playground	vineyard
bushes	farmland	horticulture	playing field	water
camping	field	marsh	plaza	wharf
woodland	fishfarm	maze	pond	

As OSM is a community-built map, there are some confusing definitions. Some of the attributes are not exclusive from the others. For example, types such as a park, garden, cemetery, etc. may contain other forms (e.g., water, grass, woodland, etc.). Therefore, to get the real picture of the destinations, OSM data are compared with the MENE survey data.

4.4.2.3 Environmental characteristics from the MENE

Table *4.5* lists the characteristics of the destinations recorded in the MENE survey and the percetages of total outdoor recreation trips made by adults from both England and North-West region. A park in a town or city is the most popular

[8] https://www.openstreetmap.org/about

destination at both national (18.2%) and regional (32%) level. A path, cycleway or bridleway possesses the next most significant share (11.36%) in England. In the North-West region, the country park claims the second place with 9.88% share. An allotment or community garden is the least popular form of destination in either England or North-West region.

Table 4.5 *Environmental Characteristics of Destinations from MENE Data*

Destinations	England %	North-West %
A beach	4.63	3.11
A children's playground	2.29	3.54
A mountain, hill or moorland	1.94	2.47
A park in a town or city	18.20	31.95
A path, cycleway or bridleway	11.36	6.74
A playing field or other recreation area	5.81	5.15
A river, lake or cannel	7.06	7.23
A village	4.91	2.77
A woodland or forest	9.60	5.40
An allotment or community garden	0.54	0.40
Another open space in a town or city	6.25	6.71
Another open space in the countryside	9.03	5.30
Country park	5.49	9.88
Farmland	6.56	3.23
Other	3.66	4.30
Other coastline	2.65	1.80

In comparing travel time across different types of destinations, there are variations among people travelling to different destinations regarding how long people travelled. As Figure 4.13 depicts, the majority travels less than an hour for outdoor recreational activities. The water feature attracts people to go for longer trips but still less than three hours. Finally, about half of the observations chose to spend their outdoor recreational time in parks, playgrounds and other city green spaces. In fact, 75% of them travelled less than 40 minutes to reach their destinations. In conclusion, the characteristics of destinations from the MENE survey is more systematic compared with the OSM data. Also, it more precisely presented the motivation of the outdoor recreational trips. Therefore, in the next section, the geographic information from the OSM and land use information from the MENE survey is combined, which will be used for model calibration.

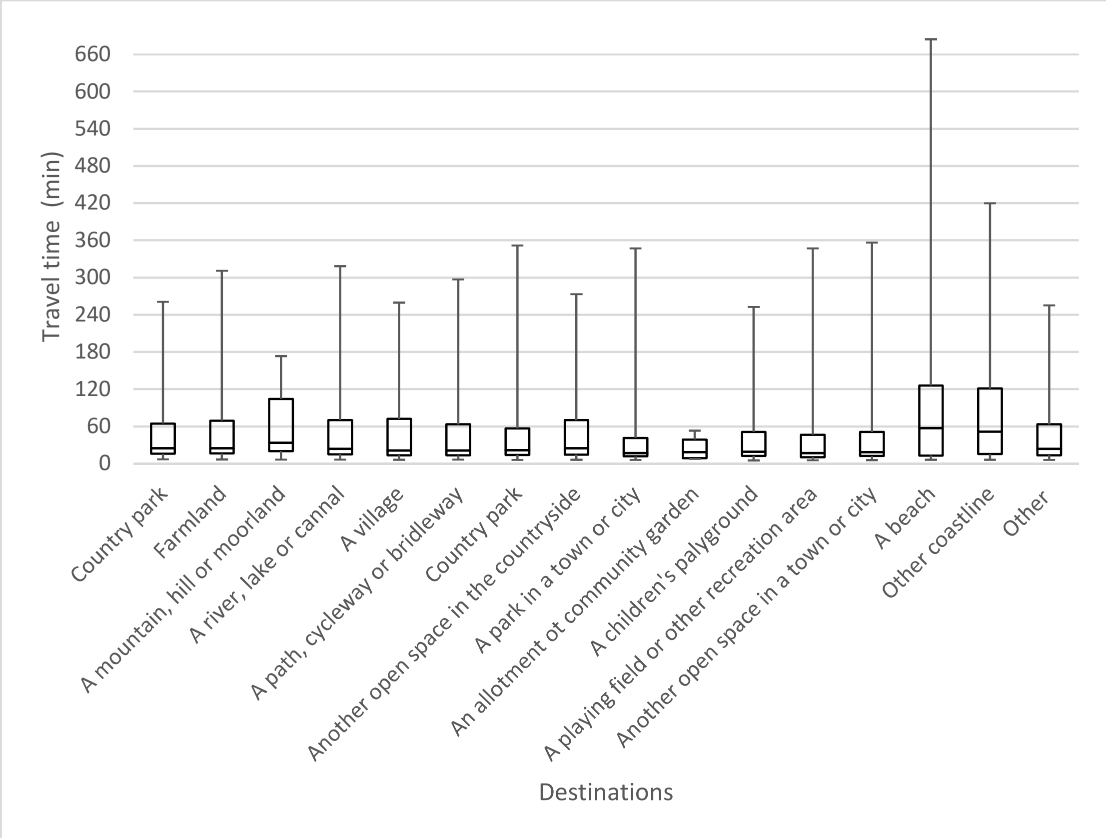


Figure 4.13. Travel time by destinations.

4.4.2.4 Combining the OSM with the MENE

As discussed above, one destination might have been identified in the MENE for several different locations which were on nearby streets: buildings. Also, the definition of a green space might be ambiguous if only refers to the OSM. Therefore, in this section, the information from the OSM and the MENE are combined in this section.

All polygons characterised by type listed in Table *4.4* from the OSM have been plotted through the Arcmap[9]. Since the classification on the OSM is fuzzy, it is only used to get geoinformation and calibrate the MENE data. The destination spots as shown in Figure 4.3 are, firstly, plotted on top of OSM polygons. In comparing the MENE environmental characteristics (Table *4.5*) with the OSM attributes (Table *4.4*), two different situations appeared:

[9] http://desktop.arcgis.com/en/arcmap/

1) The first case is the easier one, where destination points are located on an OSM polygon, and the type matches descriptions from MENE. For example, if, in the survey, people say it was ‘A park in a town or city’, the OSM map may say a park or anything that could be in a park such as water, flowerbed, recreation_ground, etc. Then we classify the polygons using MENE survey’s descriptions.
2) The second situation is more complicated. Destinations are located on an OSM polygon, but the type does not match the description, or there are no OSM polygons underneath the target points. In looking for polygons nearby (within 100 metres), if, firstly, there are any other polygons matching the description, destination points are moved; secondly, there is no polygon matching the description, insert aerial map as the base map. If it is a green space underneath the destination points, a new polygon is drawn and using the MENE information as an attribute; and finally, If it is not greenspace underneath the destination points, looking for the closed greenspace, the one match the descriptions is used; if there are more than one greenspaces matches, either the nearest one or the one with more varieties has been chosen, of course, the diversity can only be judged from the aerial map.

The result is that each destination point has a polygon underneath matching the MENE description (Figure 4.15). However, there are a few destinations whose geometric boundaries were missing, and they are under the categories around which a boundary is difficult to draw: i.e., a mountain, hill or moorland; a river, lake or canal; a village; a path, cycleway or bridleway; or a beach and other coastline. Informed by studies by Sen et al. (2014), for these destinations, the maximum area is limited to one square kilometre.

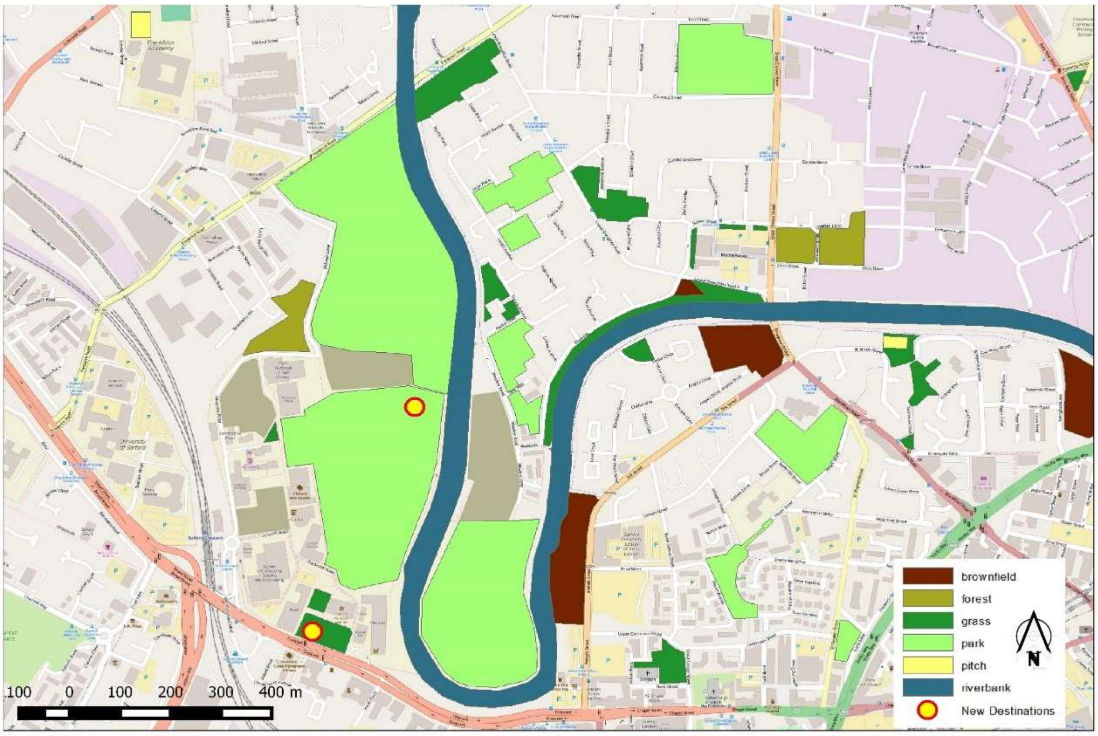


Figure 4.14. OSM 'Natural' information has been visualised through Arcmap.

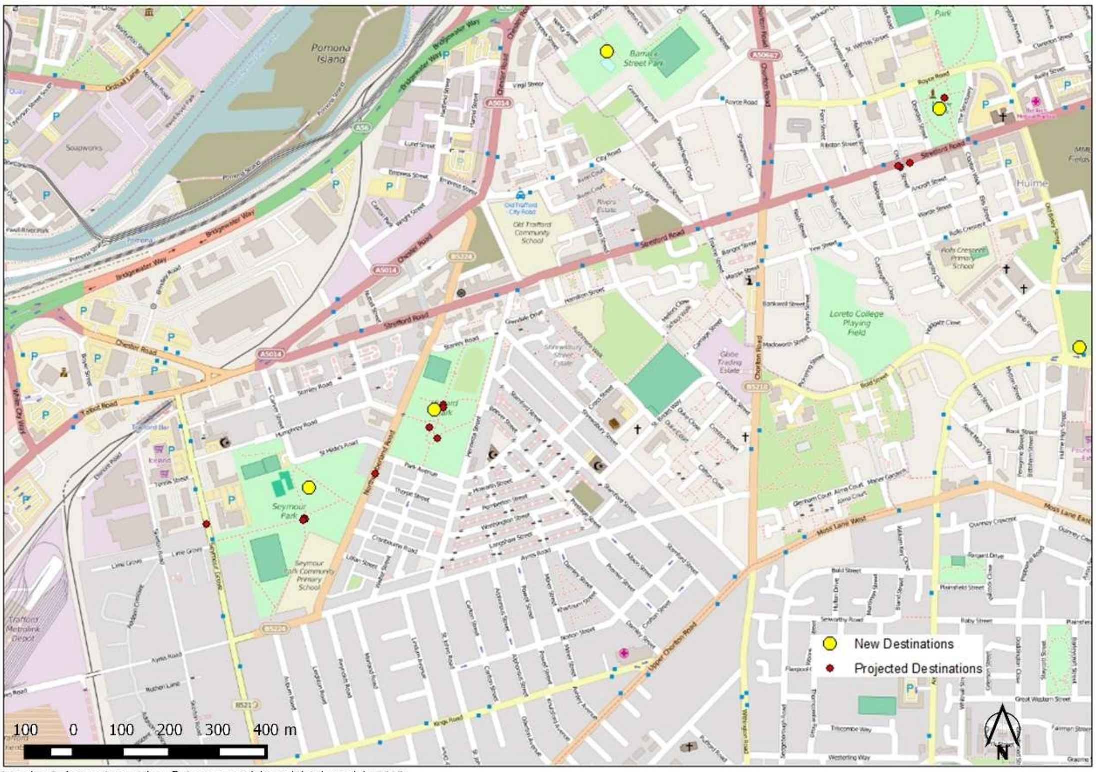


Figure 4.15. Pinpoint destination points (Yellow dots) based on observed destination projected through geocodes.

Apart from the characteristics of the outdoor recreational destinations, the land uses surrounding each destination are also tested regarding their effect on outdoor recreational demand. One of the weaknesses of OSM is that there are many voids within the OSM as it is community-built and oriented by individuals' interests. This gap is compensated by information from the MENE survey wherever it exists. However, there are still areas with missing information. Therefore, another piece of the database was involved in order to study the situation around each origin: the Generalised Land Use Database (GLUD).

4.4.2.5 The Generalised Land Use Database (GLUD)

The Generalized Land Use Database is created via a digital process, which identifies different land parcels and buildings on an Ordnance Survey digital map product, and records their 'type' and area. In brief, a classification has been developed which allocates all identifiable land features on the Ordnance Survey MasterMap into eight simplified land categories (domestic buildings, non-domestic buildings, rail, road, path, greenspace, domestic garden, water) and an additional 'unclassified' category[10]. This is too broadly classified for studying outdoor recreational destination because it does not contain details of green space characteristics. However, this is the best available land use data for studying the effects of land use outside the selected destinations.

The GLUD data has been used in two situations. Firstly, to investigate the effects of land use on surrounding origins: a 10km radius buffer area is used. This is aligned with the previous studies did by Sen et al., which shows that the 10km radius has given the best results. Land use information from the GLUD is collected to be used in Section 4.3.1, where the effects of land use on trip generation is tested. Secondly, it has been used to understand land use around every single destination; here, a 560-metre radius circle (around one square kilometre area) is applied, this is the similar with what has been applied in the UK NEA's study, where one square kilometre grid has been used to represent the site. An example is shown in Figure 4.16.

[10] https://data.gov.uk/dataset/land_use_statistics_generalised_land_use_database

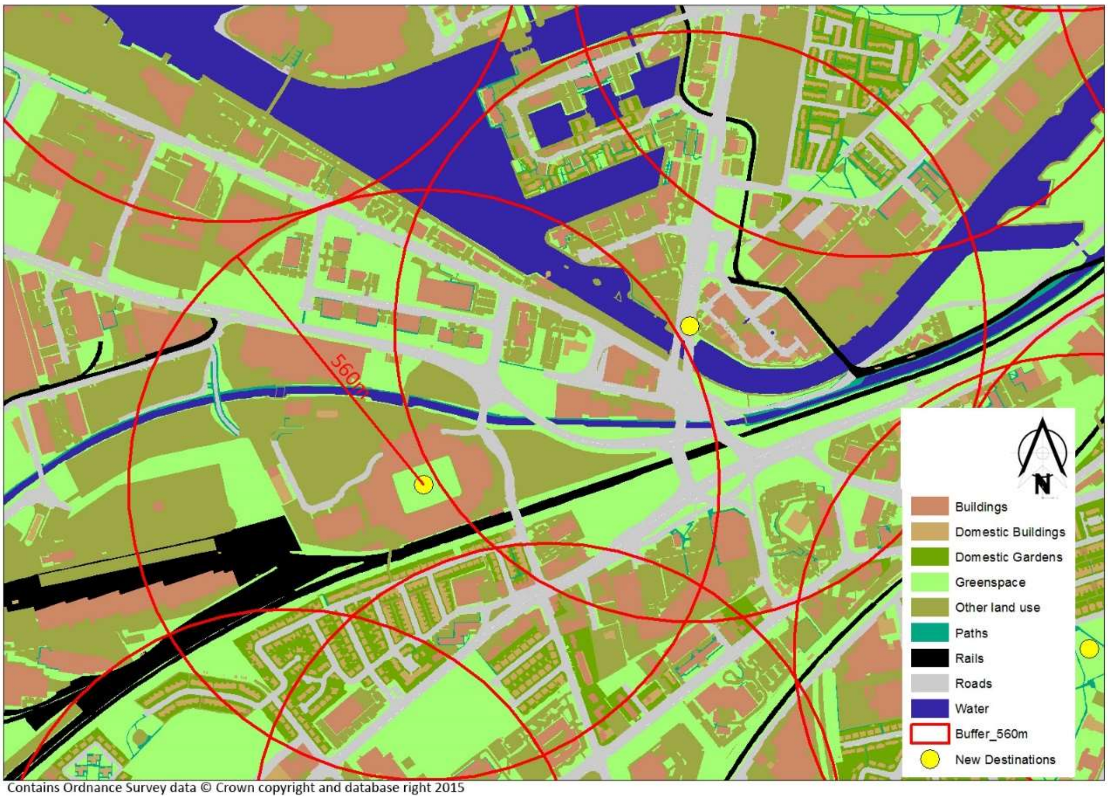


Figure 4.16. GLUD data and the 560m buffer areas on the top.

4.4.3 Demography

As shown in Table 4.6, within the study area, Manchester city has the highest population in 503,127 with 43.5 persons per hectare. The rest of the field in Greater Manchester contains a similar density around 20 persons per hectare, but districts in the Cheshire are much less crowded; in the Cheshire East, Cheshire West and Chester, the densities are 3.2 and 3.6 persons per hectare respectively. The age structures are not very different crossing the study area. About 10% of the population are between 16 to 24 years of age, about 30% are between 25 to 44 years of age, and another 25% are between 45 to 64 years of age, which leaves about 20% over 65 years of age and 15% under 16 years of age. Manchester city has the biggest ethnic minority population, about one-third of their population. This number decreases to around 15% in four adjacent districts: Bolton, Burry, Salford, and Trafford. The rest of the study area is dominated by white group.

Table 4.6 *Population and Density in Study Area*

Local Authority	Population	Density (number of persons per hectare)	Age ranges %				Non-white ethnic groups %
			16-24	25-44	45-64	65+	
Cheshire							
Cheshire - East	370,127	3.2	9.8	24.6	28.6	19.3	3.3
Cheshire-West & Chester	329,608	3.6	10.8	25.1	27.9	18.5	2.6
Halton	125,746	15.9	11.5	26.4	27.6	14.7	2.2
Warrington	202,228	11.2	10.9	27.1	26.9	15.9	4.1
Wirral	319,783	20.4	10.6	24.1	27.8	19.1	3.0
Greater Manchester							
Bolton	276,786	19.8	11.7	27.1	25.2	15.4	18.1
Bury	185,060	18.6	10.7	26.8	26.5	16.0	10.8
Manchester	503,127	43.5	19.8	33.4	18.0	9.5	33.4
Salford	233,933	24.1	13.6	29.8	23.1	14.2	9.9
Trafford	226,578	21.4	9.9	27.8	26.1	16.0	14.5
Wigan	317,849	16.9	11.0	27.2	26.8	16.3	2.7

Note: based on the census data from ONS

Following the recent study from the UK NEA (Sen et al., 2014), this research has selected four demographic variables (see Table *4.7*): percentage of retired people, non-white ethnic group, median of income, the population. As the table shows, the distributions of all four variables are skewed to the right. Median of income and non-white percentage are more significant than the other two. Therefore, they were used in their log form when investigating their correlation to the total number of trips in the Section 4.3.1.

Table 4.7 *Statistical Summary of Demographic Variables*

	Mean	Median	Standard deviation	Max	Min
Population	1,210	1,152	257	3,945	749
Retired %	13.88	13.35	5.48	34.10	0.00
Median of income £/year	26,942	24,134	10,066	73,542	9,324
Non-white %	13.53	5.00	18.51	90.10	0.30

4.4.4 Activities

One of the MENE survey questions asks for the type of recreation activities individuals took during their trips (Question 4: Which of these activities, if any, did you undertake?). Table *4.8* shows the distribution of the outdoor recreational trips in the North-West region, indicating slightly different characteristics compared with trends in England. The biggest share (about 42%) included people who went walking with dogs in England; however, this percentage in the North-West region was 21%, less than people who went for walking without a dog (28.5%). At both levels, about 60% people chose to go out for a walk. The third-favourite activity is playing with children. The difference is that, in the North-West region, the 15.53% share is twice as much as what it is at the national level.

Table 4.8 *Outdoor Recreational Activities*

Activities	England %	North-West %
Q4_Eating or Drinking out	5.08	6.55
Q4_Fieldsports (for example, shooting and hunting)	0.37	0.61
Q4_Fishing	0.49	0.80
Q4_Horsriding	0.98	0.58
Q4_Off-road cycling or mountain biking	0.91	1.19
Q4_Off-road driving or motorcycling	0.16	0.17
Q4_Picnicking	1.53	2.76
Q4_Playing with children	7.17	15.53
Q4_Road Cycling	1.83	1.82
Q4_Running	2.59	2.13
Q4_Appreciating scenery from your car	1.51	2.43
Q4_Swimming outdoors	0.41	0.41
Q4_Visits to a beach, sunbathing or padding in the sea	1.51	1.11
Q4_Visiting an attraction	2.94	3.79
Q4_Walking, not with a dog	22.39	28.49
Q4_Walking, with a dog	41.92	21.00
Q4_Watersports	0.41	0.28
Q4_Wildlife watching	2.33	2.74
Q4_Informal games and sport (for example, Frisbee or golf)	2.26	3.09
Q4_Any other outdoor activities (for example, climbing)	1.52	1.66
Q4_Others	1.68	2.87

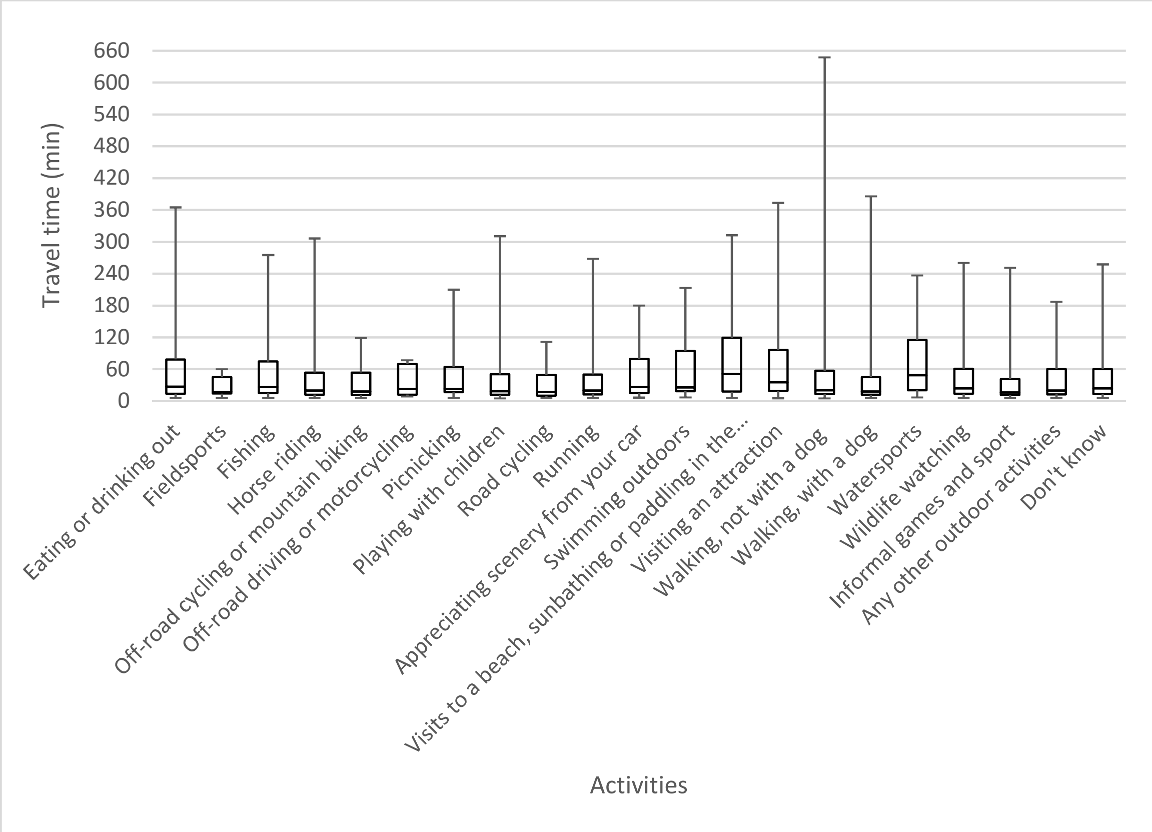


Figure 4.17. Travel time by activity.

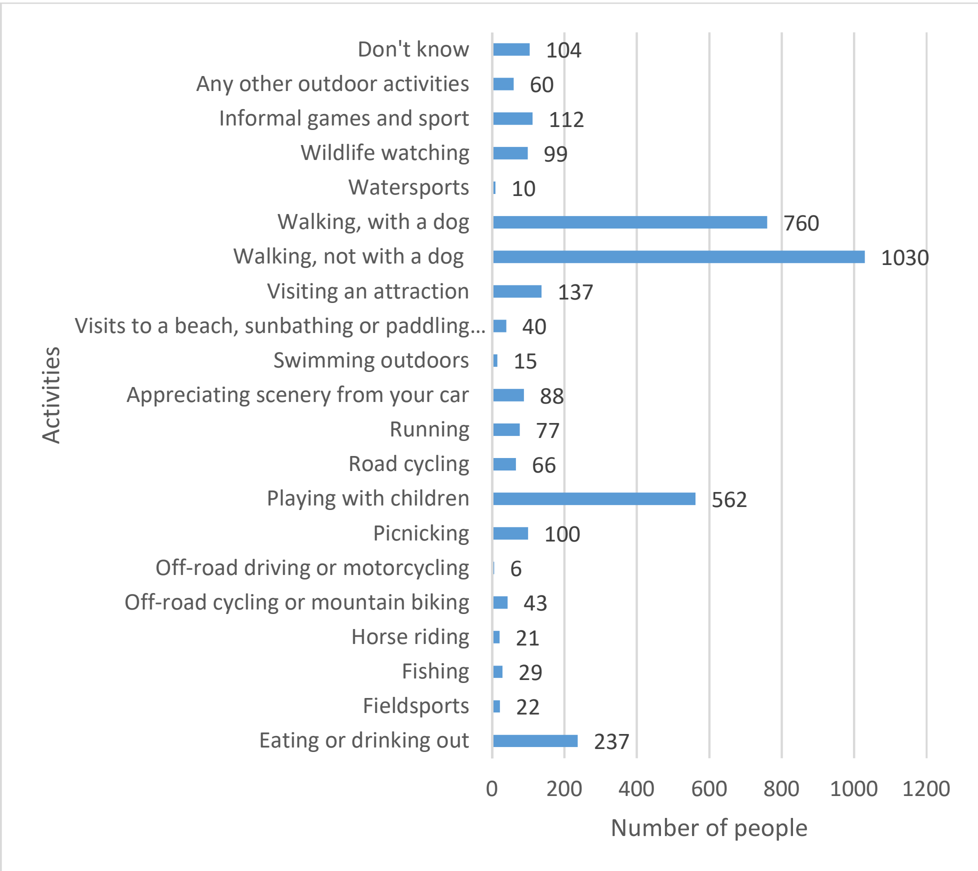


Figure 4.18. Accounts of visits by activity in study areas.

In comparing activities with travel time, as Figure 4.17 shows, 75% of observers spend about an hour on traveling to their selected destinations. Visiting an attraction, including the coast, takes longer than the others but still less than two hours for most people. Individuals who went out for sports tend to choose closer green spaces (less than 30 minutes). Based on the number of available observations in our sample area (Figure 4.18) and travel time patterns (Figure 4.17), all activities were regrouped into four new groups; Table 4.9 shows model training:

Table 4.9 *Activities Regrouping*

New activity group (Number of observations)	Activities in the MENE report
Walking, not with a dog (1030)	Walking, not with a dog
Walking with a dog (760)	Walking with a dog
Informal sports and Play (839)	Field sports, Playing with children, Running, Informal games and sport, Road cycling
Others (989)	Eating or drinking out, Fishing, Horse riding, Off-road cycling or mountain biking, Off-road driving or motorcycling, Picnicking, Appreciating scenery from your car, Swimming outdoors, Visiting beach, sunbathing or paddling in the sea, Visiting an attraction, Water sports, Wildlife watching, Any other outdoor activities

Note: Within the brackets is number of observations

4.4.5 Statistical summary of variables

Before starting the empirical experiments on model training, the variables used are summarised statistically in this section. Following the available literature and initial analysis, the following variables are selected to be used in next stage.

First, travel time is a critical factor for valuing accessibility. Based on travel pattern depicted in Figure 4.5 - Figure 4.11, when traveling time for each origin to all the destinations is calculated through Google Direction API, a filtering condition is introduced for those that do not sound achievable. For walking and cycling trips (which means people go to the destination by walking or cycling), in particular, only destinations within 20 kilometres for walking and 40 kilometres for cycling are considered to be available.

Secondly, a group of variables is used to represent characteristics of destinations recorded in the MENE survey (Table *4.10*), when an environmental characteristic described by the variable (e.g., A woodland or forest) was found in the destination, the corresponding variable is equal to one (e.g., WOODLAND = 1); otherwise, it equals to zero. An area variable is used to value the size of each green space; any space is larger than one square kilometre is normalised to one square kilometre considering movability for individuals. Also, the GLUD data are used here to describe general land use surrounding each origin and destination (Table *4.11*).

Table 4.10 *Dummy Variables Based on the MENE Survey and OSM*

Variable	Description	YES
WOODLAND	A woodland or forest (including community woodland	109
FARMLAND	Farmland or destinations locate on anything related to agriculture in the OpenStreetMap.	97
MOUNTAIN	A mountain, hill or moorland	60
WATER	A river, lake or canal	163
VILLAGE	A village	89
PATHS	A path, cycleway or bridleway	170
COUNTRYPARK	A country park	150
PARKINCITY	A park in a town or city or destinations locates on anything related to park on the OpenStreetMap.	373
ALLOTMENT	An allotment	13
PLAYGROUND	A children's playground	97
PLAYFIED	A playing field or other recreation area or destinations located on anything related to sports pitches on the OpenStreetMap.	148
IFGREEN	Any other green spaces in and around town and city	353
BEACHNCOAST	A beach and Other coastline	63

Table 4.11 *Travel Time and Area Variables*

Variable	Description	mean	Std. dev	median	Range
TIME	Traveling time from Origins to Destination in minutes	27.67	36.72	15.83	0.10-510
BUILDINGS	Coverage of non-domestic buildings	0.03	0.05	0.02	0.00-0.41
DBUILDINGS	Coverage of domestic buildings	0.05	0.04	0.05	0.00-0.20
DGARDEN	Coverage of Domestic gardens	0.15	0.12	0.14	0.00–0.61
GREENSPACES	Coverage of green spaces	0.54	0.27	0.51	0.02–1.00
ROADS	Coverage of roads	0.09	0..07	0.09	0.00-0.32
RAILS	Coverage of rail	0.01	0.01	0.00	0.00-0.12
PATHS	Coverage of path	0.01	0.01	0.01	0.00-0.07
WATER	Coverage of Water	0.05	0.10	0.01	0.00–0.88
AREA	Area of destinations based on OpenStreetMap in square kilometres. Area of those is too broad to draw a boundary (river bank, mountain, beach, etc.) are topped at 1 square kilometre.	0.50	0.44	0.35	0.00-1.00

Note: All variables start with G are derived based on the GLUD data

4.5 Trip-generation function

As stated in Chapter 3, the standard transport forecasting modelling structure is formed by stages, including trip generation, trip distribution and modal choice or the other way around (the order of trip distribution and modal choice needs to be decided through experiments). This section starts from finding out the best model form for the trip-generation function. A log linear regression model is firstly tested, followed by the category model.

4.5.1 Log linear regression analysis

This research first ran a log linear regression to see a correlation between a number of trips generated by each origin zone (LOSA) and its demographic: land use explanatory variables. The result is shown below.

Table 4.12 *Log Linear Regression Results*

Name	Value	Std err	t-test	p-value	
Intercept	-19.00	9.98	-1.91	0.06	.
Population	0.00	0.00	3.70	0.00	***
Retired	0.00	0.00	-0.66	0.51	
Income	-0.26	0.15	1.79	0.07	.
Non-white	0.00	0.00	-1.42	0.16	
Water	0.20	0.10	2.05	0.04	*
Domestic buildings	-0.19	0.13	-1.39	0.16	
Nondomestic buildings	0.39	0.29	1.35	0.18	
Roads	0.38	0.14	2.79	0.01	**
Paths	0.38	0.70	0.55	0.59	
Rails	0.25	0.34	0.73	0.47	
Greenspaces	0.20	0.10	2.04	0.04	*
Domestic Garden	0.25	0.13	1.97	0.05	*

Significant. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1
Residual standard error: 0.476 on 727 degrees of freedom
Multiple R-squared: 0.06786, Adjusted R-squared: 0.05247

Two points can be learned from the results of running the regression model as shown in Table *4.12*. First, according to the R square[11], the log linear regression does not have enough power to make robust predictions on the total number of trips generated by each origin. Second, population, accessibility (in the form of roads) and area of all types of the outdoor sites (i.e., water, greenspaces, domestic garden) are related to the number of trips generated by each LOSA area. Correlations between the total number of trips with any other variables (e.g., income, retired %, area of buildings) are not significant.

4.5.2 Trip-generation function

The log linear regression analysis indicates it is not a robust way to estimate the total number of trips from each origin. Therefore, as reviewed in Section 2.2.1, the category model should be the better model to apply. The personal category trip generation function typically divides the whole population by demographic variable. This research, as shown above, has not found trip generation to be significantly correlated to any demographic variables as revealed in Table *4.12*. Therefore, the

[11] https://en.wikipedia.org/wiki/Coefficient_of_determination

method for trip generation used here is to calculate the mean of trips per person per year based on the regional data from the MENE survey. Then it is multiplied by the population of each origin; in this case, the LOSA estimates the total number of outdoor recreational trips generated from each LSOA. This can be written as:

$$a_n = T_i / Pop_i \qquad \text{Equation 22}$$

Where a_n denotes the mean of trips made by person *n* during a year, from region *i* where individual *n* lives, T_i is the total number of trips in region *i* as recorded in the MENE data, and Pop_i is the population of region *i* according to the 2011 census data. Then, the total number of trips made from origin zone *j* with population Pop_j is written as:

$$T_j = a_n \times Pop_j \qquad \text{Equation 23}$$

Where the zone *j* is a LSOA neighbourhood, and it has to be inside the region *i* which was used to calculate a_n ; in this study, it is the North-West region.

4.6 Trip distribution and modal choice

For trip distribution and modal choice, the RUM theory-based DMCs will be used to investigate people's behaviour towards outdoor recreation. According to the reviews in Chapter 2, a RUM-based DCM is the most robust model form to study choice behaviors. Thus, the questions needed to be answered are, firstly, which form of DCMs is the best one for the outdoor recreational trips? And, secondly, what is the right order between trips distribution and modal choice? The calibration process is shown in Figure 4.1.

4.6.1 Define the model structure

4.6.1.1 Estimated standard logit model

The calibration starts with training a standard multinomial logit model; this model includes all the travel and land uses variables mentioned in Section 4.2.7, and the function takes the form of the standard logit model:

$$P_{ni} = \frac{e^{\beta\prime x_{ni}}}{\sum_j e^{\beta\prime x_{nj}}} \qquad \text{Equation 24}$$

where P_{ni} is the probability of individual *n* choosing destination *i, x* is a vector of variables relating to each destination, β is the estimated parameter for explanatory variables. The results are listed in Table 4.13.

Table 4.13 Results for a Standard Multinomial Logit Model

Name	Value	Std err	t-test	p-value	
ASC_D01	-0.169	0.173	-0.98	0.33	*[12]
ASC_D02	-0.281	0.179	-1.57	0.12	*
ASC_D03	-0.226	0.173	-1.30	0.19	*
ASC_D04	-0.0388	0.173	-0.22	0.82	*
ASC_D05	-0.306	0.18	-1.70	0.09	*
ASC_D06	-0.297	0.178	-1.67	0.10	*
ASC_D07	0.0273	0.173	0.16	0.88	*
ASC_D08	-0.207	0.174	-1.19	0.23	*
ASC_D09	-1.07	0.224	-4.78	0.00	
ASC_D10	-0.13	0.181	-0.72	0.47	*
ASC_D11	0.278	0.16	1.74	0.08	*
ASC_D12	-0.277	0.176	-1.57	0.12	*
ASC_D13	-0.248	0.181	-1.37	0.17	*
ASC_D14	-0.175	0.175	-1.00	0.32	*
ASC_D15	0.0667	0.173	0.38	0.70	*
ASC_D16	-0.0361	0.178	-0.20	0.84	*
ASC_D17	-0.626	0.188	-3.33	0.00	
ASC_D18	-0.251	0.17	-1.48	0.14	*
ASC_D19	0.049	0.169	0.29	0.77	*
ASC_D20	0	fixed			
B_TIME	-0.0648	0.00165	-39.26	0.00	
G_BUILDINGS	6.7	1.6	4.18	0.00	
G_DBUILDINGS	3.87	2.36	1.64	0.10	*
G_DGARDEN	0.513	0.945	0.54	0.59	*
G_GREENSPACE	2.51	0.805	3.12	0.00	
G_GWATER	5.09	0.888	5.73	0.00	
G_PATHS	10	1.26E-08	7.94E+08	0.00	
G_RAILS	-4.12	3.59	-1.15	0.25	*
G_ROADS	-2.85	1.66	-1.72	0.09	*
O_ALLOTMENT	-0.418	0.188	-2.22	0.03	
O_BEACHNCOAST	1.37	0.131	10.43	0.00	
O_CORRIDOR	0.146	0.0731	2.00	0.05	
O_COUNTRYPARK	0.646	0.0799	8.08	0.00	
O_FARMLAND	0.125	0.0978	1.28	0.20	*
O_IFGREEN	0.242	0.0627	3.86	0.00	
O_MOUTAIN	0.871	0.127	6.86	0.00	
O_PARKINCITY	0.454	0.0689	6.59	0.00	

[12] A sign * is appended if the absolute value of t-test is less than 1.96.

O_PLAYFIELD	0.243	0.074	3.29	0.00	
O_PLAYGROUND	0.454	0.0813	5.58	0.00	
O_VILLAGE	0.531	0.0988	5.37	0.00	
O_WATER	0.199	0.0797	2.50	0.01	
O_WOODLAND	-0.136	0.0942	-1.44	0.15	*
Number of estimated parameters:	41				
Number of observations:	2401				
Init log-likelihood:	-7192.75				
Final log-likelihood:	-2903.91				
Likelihood ratio test:	8577.696				
Rho-square:	0.596				
Adjusted rho-square:	0.591				
Diagnostic:	Convergence reached...				
Iterations:	15				
Significant. codes: '*' 0.05					

The model is estimated via software called Biogeme (Bierlaire, 2003), and all parameters starting with ASC (e.g., ASC_D01) are constants for each site. Parameters beginning with O (E.G., O_PLAYFIELD) indicate variables derived from the OpenSteetMap (OSM), and those beginning with G (e.g., G_ROADS) are variables extracted from the Generalised Land Use Database (GLUD). B_TIME is the estimated parameter for travel time, and the same definitions have been applied throughout this dissertation.

The report table also presents information of goodness of fit. The likelihood ratio index is used to measure how well the DCMs fit the data. The probability of each person in the sample choosing the alternative that was actually observed can be written as:

$$L(\beta) = \prod_{n=1}^{N} \prod_{i} (P_{ni})^{y_{ni}}$$ Equation 25

Where β is a vector containing the parameters of the model. P_{ni} is the probability of person *n* choosing the alternative *i*, y_{ni} = 1 if person *n* chose *i* and 0 otherwise. The log likelihood function is Log of $L\beta = \prod_{n=1}^{N} \prod_{i} (P_{ni})^{y_{ni}}$ Equation 25:

$$LL(\beta) = \sum_{N=1}^{N} \sum_{I} y_{ni} ln P_{ni}$$ Equation 26

The number reported in Table 4.13 is the value of β that maximises this function. Init log-likelihood is the log likelihood of the sample for the model with the parameters are set to 0. Final log-likelihood is the log likelihood of the sample for the estimated model. The likelihood ratio is calculated as:

$$-2\big(LL(0) - LL(\beta)\big) \qquad \text{Equation 27}$$

and Rho-square is calculated as:

$$\rho = 1 - \frac{LL(\widehat{\beta})}{LL(0)} \qquad \text{Equation 28}$$

Adjusted Rho-square is:

$$\rho^2 = 1 - \frac{LL(\widehat{\beta}) - K}{LL(0)} \qquad \text{Equation 29}$$

K is the number of estimated parameters. Adjusted Rho-square is the most important number to indicate how well the model fits with data. The diagnostic is the diagnostic reported by the calculation process. If the algorithm has not converged, the estimation results presented in the file cannot be trusted. Iteration is the number of iterations used by the algorithm before it stopped.

As Table 4.13 shows, a few attributes act in positive roles to attract people to use outdoor green spaces. These characteristics include ‘allotment’, ‘beach and coast’, ‘corridor’, ‘country parks’, ‘informal green space’, ‘mountain’, ‘parks in city’, ‘playfield’, ‘playground’, ‘village’ and open spaces which include water features. A reasonable weight is given to travel time (-0.06) compared with studies in any other travel demand modelling (Jin, 2002). However, some results do not align with what previous studies demonstrated. For example, woodland discourages people from visiting green spaces, and parks in the city are less attractive than village scenes. All these uncertainties lead to a question: does the main constraint of standard logit model - IIA hold in this case? In other words, is there any correlation between random parts of alternatives? Therefore, as the next step, a nested logit model is estimated on the same dataset.

4.6.1.2 Estimated nested logit model

The nested logit model is the most popular DCM to alleviate the IIA restriction on the multinomial logit model. Also, it can be used as a way to test the IIA in the multinomial logit model as reviewed in Section 2.3.2. By dividing options into subgroups, the IIA assumption of standard logit model is alleviated in the nested logit model, but it still needs to be held within each subset. There are two ways of dividing the alternatives that are widely used in the transportation modelling: split by transport mode or divided by destination. This research examines both forms.

Travel mode on the upper level

Firstly, the transport mode choice is placed at the top level and then destination (Figure 4.19).

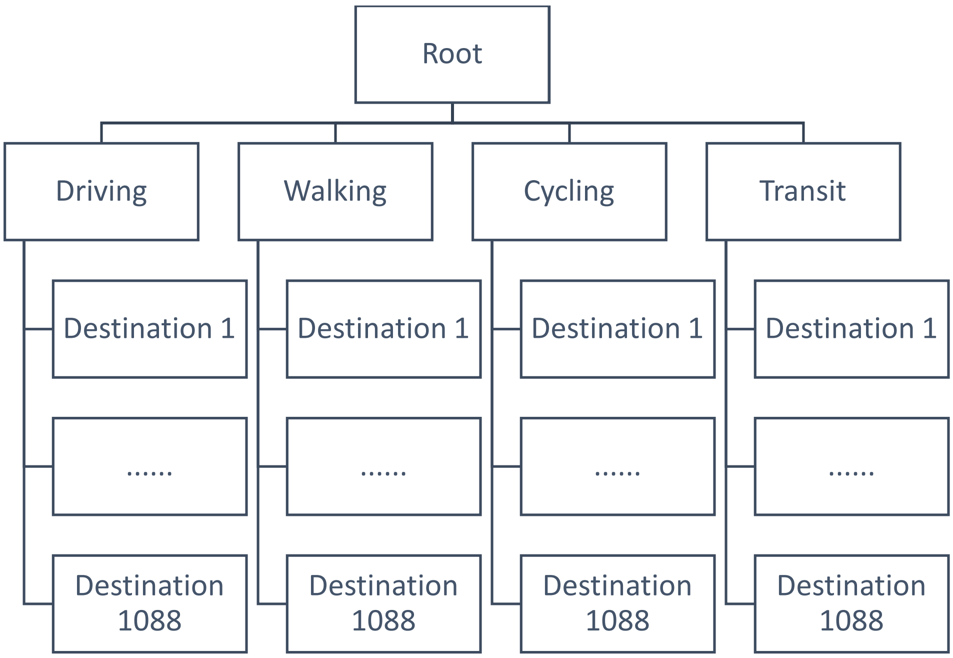


Figure 4.19. Nests structure for travel mode choice first nested logit model. "……" (Ellipsis) means Destination 2 to Destination 1087.

The results are shown in Table 4.14 and Table 4.15.

Table 4.14 *Results for Destination Choice Logit Model as Lower Part of Nested Logit Model*

Name	Value	Robust Std err	Robust t-test	p-value	
ASC_D01	0				
ASC_D02	-9.62E-05	0.123	0.00	1.00	*
ASC_D03	0.0512	0.121	0.42	0.67	*
ASC_D04	0.0149	0.122	0.12	0.9	*
ASC_D05	-0.0229	0.123	-0.19	0.85	*
ASC_D06	0.0513	0.121	0.42	0.67	*
ASC_D07	0.0512	0.121	0.42	0.67	*
ASC_D08	-7.08E-06	0.123	0.00	1.00	*
ASC_D09	-2.91E-05	0.123	0.00	1.00	*
ASC_D10	0.015	0.122	0.12	0.90	*
ASC_D11	-7.97E-05	0.123	0.00	1.00	*
ASC_D12	0.0222	0.122	0.18	0.86	*
ASC_D13	-0.0227	0.123	-0.18	0.85	*
ASC_D14	0.0149	0.122	0.12	0.90	*
ASC_D15	0.0512	0.121	0.42	0.67	*
ASC_D16	-9.74E-05	0.123	0.00	1.00	*
ASC_D17	-0.0228	0.123	-0.18	0.85	*
ASC_D18	0.0149	0.122	0.12	0.90	*
ASC_D19	-0.103	0.126	-0.82	0.41	*
ASC_D20	-0.0863	0.125	-0.69	0.49	*
O_AREA	0.00409	0.0247	0.17	0.87	*
O_BEACHNCOAST	-0.0124	0.0882	-0.14	0.89	*
O_CORRIDOR	-0.0029	0.0786	-0.04	0.97	*
O_FARMLAND	-0.00111	0.103	-0.01	0.99	*
O_IFGREEN	-0.00474	0.12	-0.04	0.97	*
O_MOUTAIN	-0.00259	0.119	-0.02	0.98	*
O_PARKINCITY	0.00377	0.0525	0.07	0.94	*
O_PLAYFIELD	-0.00022	0.055	0.00	1.00	*
O_WATER	-0.00818	0.0713	-0.11	0.91	*
O_VILLAGE	-0.00872	0.127	-0.07	0.95	*
O_WOODLAND	-0.00703	0.0614	-0.11	0.91	*
G_BUILDINGS	-0.0887	1.05	-0.08	0.93	*
G_DGARDEN	-0.048	0.496	-0.10	0.92	*
G_GREENSPACE	-0.0452	0.552	-0.08	0.93	*
G_ROADS	-0.0395	0.904	-0.04	0.97	*
G_GWATER	-0.0297	0.566	-0.05	0.96	*

Number of estimated parameters:	36
Number of observations:	2401
Null log-likelihood:	-7986.622
Init log-likelihood:	-7986.622
Final log-likelihood:	-7984.532
Likelihood ratio test:	4.181
Rho-square:	0
Adjusted rho-square:	-0.004
Diagnostic:	Convergence reached

Iterations:	2
Significant. codes: ‘*’ > 0.05	

Table 4.15 *Results for Mode Choice Logit Model as Upper Part of Nest Logit Mode*

Name	Value	Robust Std err	Robust t-test	p-value	
ASC_Car	0				
ASC_Cycling	-1.92	0.139	-13.82	0	
ASC_Transit	-1.36	0.159	-8.51	0	
ASC_Walking	1.74	0.177	9.87	0	
B_LOGSUM	4.02E-16	1.01E+07	0.00	1	*
B_TIME	-0.03	0.00498	-6.03	0	
Number of estimated parameters:	5				
Number of observations:	2401				
Null log-likelihood:	-3445.722				
Init log-likelihood:	-3445.722				
Final log-likelihood:	-1815.574				
Likelihood ratio test:	3260.297				
Rho-square:	0.473				
Adjusted rho-square:	0.472				
Diagnostic:	Convergence reached.				
Iterations:	8				
Significant. codes: ‘*’ > 0.05					

Note: B_LOGSUM is a parameter for inclusive value calculated based on results from table 4.15.

As the tables suggest, travel time is the only useful predictor (Table *4.14*) in this model. Destination attributes do not work without travel time (Table 4.15).

Destinations choice on the upper level

The second structure is the other way around (Figure 4.20). This assumes that people decide where to go first and then choose the suitable transport mode; results are shown in Table 4.16 and Table 4.17.

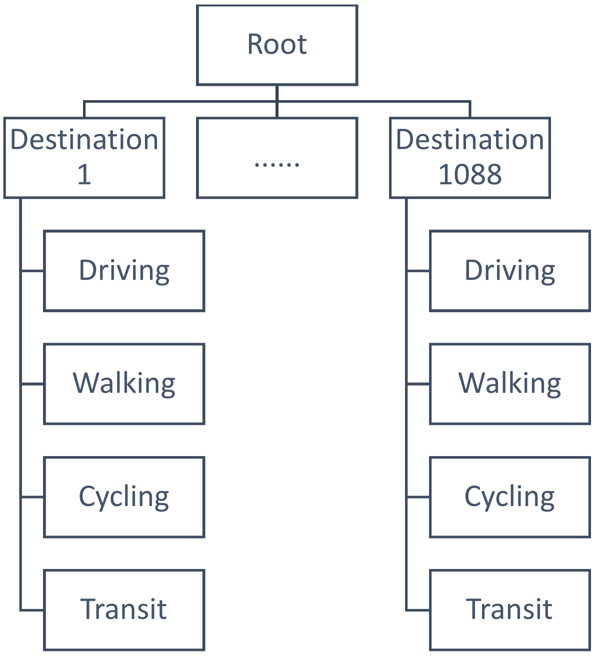


Figure 4.20. Nests structure for destination choice first nested logit model. "……" (Ellipsis) means Destination 2 to Destination 1087.

Table 4.16 *Results for Mode Choice Logit Model as Lower Part of Nested Logit Model*

Name	Value	Robust Std err	Robust t-test	p-value
ASC_Car	0			
ASC_Cycling	-1.92	0.136	-14.19	0
ASC_Transit	-1.36	0.157	-8.66	0
ASC_Walking	1.74	0.175	9.96	0
B_TIME	-0.03	0.00496	-6.05	0
Number of estimated parameters:	4			
Number of observations:	2401			
Null log-likelihood:	-3445.72			
Init log-likelihood:	-3445.72			
Final log-likelihood:	-1815.57			
Likelihood ratio test:	3260.297			
Rho-square:	0.473			
Adjusted rho-square:	0.472			
Diagnostic:	Convergence reached...			
Iterations:	8			

Table 4.17 *Results for Destination Choice Logit Mode as Upper Part of Nested Logit Model*

Name	Value	Robust Std err	Robust t-test	p-value	
ASC_D01	0				
ASC_D02	0.0614	0.14	0.44	0.66	*
ASC_D03	0.0491	0.133	0.37	0.71	*
ASC_D04	0.0413	0.135	0.31	0.76	*
ASC_D05	-0.0394	0.14	-0.28	0.78	*
ASC_D06	0.0175	0.135	0.13	0.90	*
ASC_D07	0.076	0.135	0.56	0.57	*
ASC_D08	-0.0617	0.138	-0.45	0.65	*
ASC_D09	0.048	0.136	0.35	0.72	*
ASC_D10	-0.057	0.135	-0.42	0.67	*
ASC_D11	-0.0527	0.138	-0.38	0.70	*
ASC_D12	-0.0377	0.139	-0.27	0.79	*
ASC_D13	0.0141	0.136	0.10	0.92	*
ASC_D14	-0.0215	0.137	-0.16	0.88	*
ASC_D15	0.0759	0.139	0.55	0.59	*
ASC_D16	0.125	0.135	0.92	0.36	*
ASC_D17	0.038	0.139	0.27	0.78	*
ASC_D18	-0.00453	0.132	-0.03	0.97	*
ASC_D19	-0.0984	0.142	-0.69	0.49	*
ASC_D20	0.0124	0.14	0.09	0.93	*
B_LOGSUM	-0.0517	0.00118	-43.72	0.00	
G_BUILDINGS	5.59	1.32	4.24	0.00	
G_DGARDEN	0.225	0.6	0.38	0.71	*
G_GREENSPACE	2	0.675	2.96	0.00	
G_GWATER	4.14	0.714	5.80	0.00	
G_ROADS	-2.98	1.03	-2.88	0.00	
O_AREA	0.0695	0.0321	2.16	0.03	
O_BEACHNCOAST	0.581	0.112	5.20	0.00	
O_CORRIDOR	-0.188	0.0978	-1.93	0.05	*
O_FARMLAND	-0.00916	0.12	-0.08	0.94	*
O_IFGREEN	0.0019	0.143	0.01	0.99	*
O_MOUTAIN	0.844	0.143	5.89	0.00	
O_PARKINCITY	-0.245	0.0639	-3.83	0.00	
O_PLAYFIELD	-0.236	0.0592	-3.99	0.00	
O_VILLAGE	0.609	0.171	3.56	0.00	
O_WATER	-0.0238	0.0866	-0.27	0.78	*
O_WOODLAND	0.148	0.0687	2.16	0.03	

Number of estimated parameters:	37
Number of observations:	2401
Null log-likelihood:	-7986.62
Init log-likelihood:	-7986.62
Final log-likelihood:	-4604.5
Likelihood ratio test:	6764.238
Rho-square:	0.423
Adjusted rho-square:	0.419

Diagnostic:	Convergence reached...
Iterations:	14
Significant. codes: '*' > 0.05	

This model structure is not applicable if we compare the parameter for travel time (B_TIME) in Table 4.16 with the inclusive value of transport modes (B_LOGSUM) in Table 4.17. This is because the absolute value of B_TIME (0.03) is smaller than its upper level form B_LOGSUM (0.05). As William (1977) stated, parameters of the multinomial logit models associated with different level should not increase under progression towards the top.

Based on the above experiments, simply using a nested logit model is apparently not applicable in this case. However, we can learn some information from this experiment which is useful for further developing the model form. Firstly, the alternatives are correlated by transport mode, which means the options have to be regrouped by transport mode before they can be used in the multinomial logit model. Secondly, in outdoor recreational trips, the modal choice comes before destination choice. This finding is aligned with the results from Rohr (2005). Finally, the model form for outdoor recreational trips has to include land use and environmental attributes as well as travel time. Therefore, instead of modelling the outdoor recreational trips altogether, the samples will be grouped by transport mode, and the multinomial logit model will be run, including both travel time and land uses variables on the subsets. Before doing this, another popular form of DCM—the mixed logit model—is tested in the next section to investigate the variations among individuals for each explanatory variable.

4.6.1.3 Estimated mixed logit model

The mixed logit model is tested in this section to investigate the variation of each explanatory variable among individuals. The mixed logit model was developed by Train (2009). The most common form of mixed logit model is the random parameters logit model. Because the parameters are replaced with random coefficients which incorporate individual heterogeneity, the random component of each parameter will be shared across choice alternatives, and very general patterns of correlation will be allowed in this model. As described in Section 2.3.3, the utility function of the random coefficients model is written as:

$$U_{nj} = \beta_n' x_{nj} + \varepsilon_{nj} \qquad \text{Equation 30}$$

where x_{nj} are observed variables, β_n' is a vector of coefficients of these variables for person *n* representing that person's tastes, and ε_{nj} is a random term. The only difference between random coefficients-mixed logit model and standard logit model is the coefficients vary over decision makers in the population with density $f(\beta)$. In this model, $f(\beta)$ is specified to be normal distributed. Results are shown in Table 4.18.

Table 4.18 *Results of Mixed Logit (Random Parameter) Model*

Name	Value	Std err	t-test	p-value	
ASC_D01	-0.167	0.229	-0.73	0.47	*
ASC_D02	-0.363	0.239	-1.51	0.13	*
ASC_D03	-0.293	0.233	-1.26	0.21	*
ASC_D04	-0.13	0.232	-0.56	0.58	*
ASC_D05	-0.285	0.239	-1.19	0.23	*
ASC_D06	-0.332	0.240	-1.38	0.17	*
ASC_D07	0.0515	0.232	0.22	0.82	*
ASC_D08	-0.232	0.235	-0.99	0.32	*
ASC_D09	-1.28	0.293	-4.38	0.00	
ASC_D10	-0.233	0.241	-0.97	0.33	*
ASC_D11	0.489	0.212	2.30	0.02	
ASC_D12	-0.282	0.235	-1.20	0.23	*
ASC_D13	-0.22	0.242	-0.91	0.36	*
ASC_D14	-0.085	0.228	-0.37	0.71	*
ASC_D15	-0.0116	0.235	-0.05	0.96	*
ASC_D16	-0.0305	0.237	-0.13	0.90	*
ASC_D17	-0.793	0.260	-3.05	0.00	
ASC_D18	-0.246	0.227	-1.09	0.28	*
ASC_D19	0.0688	0.225	0.31	0.76	*
ASC_D20	0	fixed			
B_TIME	-0.155	0.00815	-19.01	0.00	
B_TIME_S	-0.162	0.00999	-16.25	0.00	
G_BUILDINGS	7.23	2.25	3.21	0.00	
G_BUILDINGS_S	-1.01	10.8	-0.09	0.93	*
G_DBUILDINGS	0.69	3.29	0.21	0.83	*
G_DBUILDINGS_S	0.0203	6.7	0.00	1.00	*
G_DGARDEN	-0.0443	1.27	-0.03	0.97	*
G_DGARDEN_S	-0.488	3.66	-0.13	0.89	*
G_GREENSPACE	1.98	1.09	1.82	0.07	*
G_GREENSPACE_S	-5.12	0.82	-6.24	0.00	

G_GWATER	4.2	1.26	3.34	0.00	
G_GWATER_S	-7.92	2.25	-3.53	0.00	
G_PATHS	10	4.52E-08	2.21E+08	0.00	
G_PATHS_S	-0.0244	1.80e+308	0.00	1.00	*
G_RAILS	-9.63	4.81	-2.00	0.05	
G_RAILS_S	-0.154	1.80e+308	0.00	1.00	*
G_ROADS	-6.36	2.27	-2.80	0.01	
G_ROADS_S	-0.252	3.72	-0.07	0.95	*
O_ALLOTMENT	-0.554	0.263	-2.11	0.04	
O_ALLOTMENT_S	0.121	1.43	0.08	0.93	*
O_BEACHNCOAST	1.77	0.204	8.68	0.00	
O_BEACHNCOAST_S	-2.19	0.863	-2.53	0.01	
O_CORRIDOR	0.166	0.0995	1.67	0.10	*
O_CORRIDOR_S	-0.0298	0.48	-0.06	0.95	*
O_COUNTRYPARK	0.845	0.108	7.84	0.00	
O_COUNTRYPARK_S	0.0474	0.515	0.09	0.93	*
O_FARMLAND	0.15	0.133	1.12	0.26	*
O_FARMLAND_S	0.105	0.731	0.14	0.89	*
O_IFGREEN	0.321	0.0847	3.79	0.00	
O_IFGREEN_S	0.23	0.832	0.28	0.78	*
O_MOUTAIN	0.886	0.218	4.07	0.00	
O_MOUTAIN_S	2.49	0.689	3.61	0.00	
O_PARKINCITY	0.591	0.0942	6.27	0.00	
O_PARKINCITY_S	0.0118	0.514	0.02	0.98	*
O_PLAYFIELD	0.318	0.101	3.15	0.00	
O_PLAYFIELD_S	-0.00388	0.56	-0.01	0.99	*
O_PLAYGROUND	0.469	0.113	4.17	0.00	
O_PLAYGROUND_S	0.00302	0.474	0.01	0.99	*
O_VILLAGE	0.738	0.187	3.95	0.00	
O_VILLAGE_S	-0.312	5.06	-0.06	0.95	*
O_WATER	0.309	0.107	2.88	0.00	
O_WATER_S	-0.114	0.667	-0.17	0.86	*
O_WOODLAND	-0.179	0.13	-1.38	0.17	*
O_WOODLAND_S	0.098	0.61	0.16	0.87	*

Number of draws:	5000
Number of estimated parameters:	63
Number of observations:	2401
Null log-likelihood:	-7192.753
Init log-likelihood:	-7192.753
Final log-likelihood:	-2605.353
Likelihood ratio test:	9174.8
Rho-square:	0.638
Adjusted rho-square:	0.629
Diagnostic:	Convergence reached...
Iterations:	263
Significant. codes: ‘*’ > 0.05	

Since every parameter is assumed to be normally distributed, the estimated results are formed by two parts: using BUILDING as an example: G_BUILDING represents the mean of the coefficients for individuals and G_BUILDING_S denotes the standard deviation.

Table *4.18* contains the variables whose means and standard deviations have both reached significant levels (*t*-test result is bigger than 1.96): travel time, coverage of water, characteristic of beach or coastline, mountain. These variables are suggested as useful predictors and have considerable variation among individuals. The significant variation in travel time needs to deal with different transport modes separately. Regarding variations on water feature and mountains, this is something that has not been studied in this research but clearly needs further investigation.

Coverage of buildings and roads, categories of allotment, country park, informal green spaces, parks, playfield, playground, village, and water are believed to be significantly related to the number of visits, and the level of impacts is very similar to everyone. The rest of the variables are considered to be unimportant regarding outdoor recreational trips.

Although the random parameter logit model has improved the model results, making predictions through this model is based on simulation, and the coefficients will be random but normally distributed. Results are difficult to be explained and applied for planning purposes. For this reason, the new model in the following sections will be developed as summarised in Section 4.3.2.2, in the form of a group of multinomial models, which will be run on the subsets grouped by transport mode.

4.6.1.4 Estimated standard logit models by different transport mode

In the following sections, alternatives were divided by transport mode and tested by the multinomial logit model. Besides changing the model structure, constants for sites (ASC) will be not be presented in the results table anymore. This is because the option pool will be too big if all the 1088 destinations are shown in Figure 4.3. The sampling method was used in this research is called the random sampling method. The options pool is formed by 20 destinations: one selected site, and 19 alternatives which are randomly drawn from all 1088 destinations. The order of these 20 options was also drawn randomly. If one constant reaches the statistical significant level consistently, it means that the destination in that place will always be favoured regardless what the destination is. In other words, the significant results produced by the model are false positive. The technical details of random sampling can be seen in Appendix A. As Table 4.13 to Table *4.18* show, constants are rarely relevant, and, even when they happen, no more than three ASCs appeared to be significant, and the constants appearing to be important are random instead of consistent. It means there is no evidence that the computer could recognise which alternative is favoured as a default. For this reason, constants are removed from following test.

Table 4.19 *Model Results for Cycling Trips*

Name	Value	Std err	t-test	p-value	
B_TIME	-0.149	0.0277	-5.36	0.00	
G_BUILDINGS	3.57	13.8	0.26	0.80	*
G_DBUILDINGS	10	1.40E-08	7.12E+08	0.00	
G_DGARDEN	-4.47	9.33	-0.48	0.63	*
G_GREENSPACE	-0.979	8.2	-0.12	0.90	*
G_GWATER	7.44	9.5	0.78	0.43	*
G_PATHS	-10	2.58E-08	-3.9E+08	0.00	
G_RAILS	1.51	29.9	0.05	0.96	*
G_ROADS	-7.54	16.8	-0.45	0.65	*
O_ALLOTMENT	-5.66	2.95	-1.92	0.05	*
O_BEACHNCOAST	-0.55	1.63	-0.34	0.74	*
O_CORRIDOR	1.77	0.678	2.62	0.01	
O_COUNTRYPARK	2.51	0.901	2.78	0.01	
O_FARMLAND	1.05	1.22	0.86	0.39	*
O_IFGREEN	0.658	0.584	1.13	0.26	*
O_MOUTAIN	5.88	1.57	3.74	0.00	
O_PARKINCITY	-0.214	0.81	-0.26	0.79	*
O_PLAYFIELD	0.517	0.73	0.71	0.48	*
O_PLAYGROUND	-0.621	1.06	-0.58	0.56	*
O_VILLAGE	0.684	1.5	0.46	0.65	*
O_WATER	1.35	0.994	1.36	0.17	*
O_WOODLAND	-0.811	1.21	-0.67	0.50	*
Number of estimated parameters:	41				
Number of observations:	71				
Null log-likelihood:	-212.697				
Init log-likelihood:	-212.697				
Final log-likelihood:	-35.707				
Likelihood ratio test:	353.98				
Rho-square:	0.832				
Adjusted rho-square:	0.639				
Diagnostic:	Convergence reached...				
Iterations:	20				
Significant. codes: ‘*’ > 0.05					

Table 4.20 *Model Results for Driving Trips*

Name	Value	Std err	t-test	p-value	
B_TIME	-0.0526	0.00238	-22.12	0.00	
G_BUILDINGS	5.8	2.05	2.84	0.00	
G_DBUILDINGS	0.814	3.46	0.24	0.81	*
G_DGARDEN	-0.231	1.23	-0.19	0.85	*
G_GREENSPACE	1.2	0.971	1.24	0.22	*
G_GWATER	3.73	1.05	3.54	0.00	
G_PATHS	10	1.80e+308	0.00	1.00	*
G_RAILS	1.36	4.44	0.31	0.76	*
G_ROADS	-4.97	2.11	-2.36	0.02	
O_ALLOTMENT	-0.391	0.218	-1.79	0.07	*
O_BEACHNCOAST	1.11	0.144	7.72	0.00	
O_CORRIDOR	-0.0646	0.0987	-0.65	0.51	*
O_COUNTRYPARK	0.739	0.0964	7.67	0.00	
O_FARMLAND	0.106	0.111	0.96	0.34	*
O_IFGREEN	0.0559	0.0795	0.70	0.48	*
O_MOUTAIN	0.828	0.141	5.86	0.00	
O_PARKINCITY	0.284	0.0894	3.18	0.00	
O_PLAYFIELD	0.299	0.0925	3.23	0.00	
O_PLAYGROUND	0.576	0.105	5.51	0.00	
O_VILLAGE	0.551	0.115	4.80	0.00	
O_WATER	0.15	0.095	1.58	0.12	*
O_WOODLAND	0.126	0.108	1.16	0.25	*
Number of estimated parameters:	41				
Number of observations:	995				
Null log-likelihood:	-2980.75				
Init log-likelihood:	-2980.75				
Final log-likelihood:	-2063.77				
Likelihood ratio test:	1833.958				
Rho-square:	0.308				
Adjusted rho-square:	0.294				
Diagnostic:	Radius of the trust region is too small				
Iterations:	45				
Significant. codes: ‘*’ > 0.05					

Table 4.21 *Model Results for Transit Trips*

Name	Value	Std err	t-test	p-value	
B_TIME	-0.0368	0.00609	-6.03	0.00	
G_BUILDINGS	0.828	5.26	0.16	0.87	*
G_DBUILDINGS	10	2.40E-08	4.17E+08	0.00	
G_DGARDEN	-8.26	3.8	-2.18	0.03	
G_GREENSPACE	-1.65	3.17	-0.52	0.60	*
G_GWATER	3.01	3.45	0.87	0.38	*
G_PATHS	10	3.94E-08	2.54E+08	0.00	
G_RAILS	10	3.17E-08	3.15E+08	0.00	
G_ROADS	-1.97	6.77	-0.29	0.77	*
O_ALLOTMENT	-1.11	0.998	-1.11	0.27	*
O_BEACHNCOAST	1.77	0.585	3.03	0.00	
O_CORRIDOR	0.589	0.412	1.43	0.15	*
O_COUNTRYPARK	0.904	0.479	1.89	0.06	*
O_FARMLAND	-0.27	0.668	-0.40	0.69	*
O_IFGREEN	0.38	0.345	1.10	0.27	*
O_MOUTAIN	0.518	1.11	0.47	0.64	*
O_PARKINCITY	0.656	0.384	1.71	0.09	*
O_PLAYFIELD	0.411	0.451	0.91	0.36	*
O_PLAYGROUND	0.851	0.404	2.10	0.04	
O_VILLAGE	1.54	0.561	2.74	0.01	
O_WATER	0.0166	0.417	0.04	0.97	*
O_WOODLAND	-0.718	0.601	-1.20	0.23	*
Number of estimated parameters:	41				
Number of observations:	84				
Null log-likelihood:	-251.642				
Init log-likelihood:	-251.642				
Final log-likelihood:	-110.086				
Likelihood ratio test:	283.11				
Rho-square:	0.563				
Adjusted rho-square:	0.4				
Diagnostic:	Convergence reached...				
Iterations:	20				
Significant. codes: ‘*’ > 0.05					

Table 4.22 *Model Results for Walking Trips*

Name	Value	Std err	t-test	p-value	
B_TIME	-0.0866	0.00394	-21.99	0.00	
G_BUILDINGS	3.2	4.29	0.74	0.46	*
G_DBUILDINGS	8.75	4.83	1.81	0.07	*
G_DGARDEN	1.46	2.02	0.72	0.47	*
G_GREENSPACE	4.72	1.94	2.44	0.01	
G_GWATER	3.14	2.24	1.40	0.16	*
G_PATHS	10	1.84E-08	5.44E+08	0.00	
G_RAILS	-10	9.28E-09	-1.1E+09	0.00	
G_ROADS	-1.31	3.8	-0.34	0.73	*
O_ALLOTMENT	0.0312	0.552	0.06	0.95	*
O_BEACHNCOAST	2.35	0.45	5.22	0.00	
O_CORRIDOR	0.514	0.159	3.24	0.00	
O_COUNTRYPARK	0.562	0.201	2.80	0.01	
O_FARMLAND	0.146	0.289	0.51	0.61	*
O_IFGREEN	0.58	0.14	4.15	0.00	
O_MOUTAIN	1.28	0.371	3.46	0.00	
O_PARKINCITY	0.963	0.153	6.28	0.00	
O_PLAYFIELD	0.223	0.173	1.29	0.20	*
O_PLAYGROUND	0.209	0.185	1.13	0.26	*
O_VILLAGE	0.352	0.262	1.34	0.18	*
O_WATER	0.785	0.216	3.64	0.00	
O_WOODLAND	-0.876	0.253	-3.46	0.00	
Number of estimated parameters:	41				
Number of observations:	1251				
Null log-likelihood:	-3747.66				
Init log-likelihood:	-3747.66				
Final log-likelihood:	-498.593				
Likelihood ratio test:	6498.136				
Rho-square:	0.867				
Adjusted rho-square:	0.856				
Diagnostic:	Convergence reached...				
Iterations:	17				
Significant. codes: '*' > 0.05					

While cycling, transit and walking models return encouraging estimations, the driving model is not working judging by the diagnostic in Table 4.20 says the estimation is not converged. In the following steps therefore, further improvements of the model will be done through adjusting the variables and alternatives.

4.6.2 Finalise exploratory variables

4.6.2.1 OpenStreetMap variables only

In this section, the GLUD variables are replaced by an area variable O_AREA which is derived from the OSM. This aims to reduce the similarities between land uses and types of outdoor recreational destinations, without sacrificing the size factor which was represented by the GLUD variables at the moment. The results are shown below.

Table 4.23 *Model Results of Cycling Trips, OSM Variables Only*

Name	Value	Std err	t-test	p-value	
B_TIME	-0.161	0.0307	-5.25	0.00	
O_ALLOTMENT	-4.22	2.23	-1.89	0.06	*
O_AREA	1.12	0.364	3.06	0.00	
O_BEACHNCOAST	1.34	1.55	0.86	0.39	*
O_CORRIDOR	1.72	0.689	2.49	0.01	
O_COUNTRYPARK	2.85	0.878	3.25	0.00	
O_FARMLAND	0.274	1.1	0.25	0.80	*
O_IFGREEN	0.872	0.618	1.41	0.16	*
O_MOUTAIN	6.02	1.59	3.80	0.00	
O_PARKINCITY	-0.458	0.756	-0.61	0.54	*
O_PLAYFIELD	0.562	0.754	0.75	0.46	*
O_PLAYGROUND	-0.543	0.998	-0.54	0.59	*
O_VILLAGE	0.489	1.26	0.39	0.70	*
O_WATER	1.97	0.897	2.19	0.03	
O_WOODLAND	-1.31	1.21	-1.08	0.28	*
Number of estimated parameters:	34				
Number of observations:	71				
Null log-likelihood:	-212.697				
Init log-likelihood:	-212.697				
Final log-likelihood:	-35.319				
Likelihood ratio test:	354.755				
Rho-square:	0.834				
Adjusted rho-square:	0.674				
Diagnostic:	Convergence reached...				
Iterations:	15				
Significant. codes: ‘*’ > 0.05					

Table 4.24 *Model Results for Driving Trips, OSM Variables Only*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.0507	0.00487	-10.42	0.00	
O_ALLOTMENT	-0.119	0.189	-0.63	0.53	*
O_AREA	0.358	0.0253	14.18	0.00	
O_BEACHNCOAST	1.73	0.138	12.54	0.00	
O_CORRIDOR	-0.305	0.0868	-3.51	0.00	
O_COUNTRYPARK	0.973	0.0869	11.20	0.00	
O_FARMLAND	0.231	0.109	2.11	0.03	
O_IFGREEN	0.205	0.0775	2.64	0.01	
O_MOUTAIN	1.01	0.145	6.96	0.00	
O_PARKINCITY	0.137	0.0802	1.70	0.09	*
O_PLAYFIELD	0.265	0.0876	3.02	0.00	
O_PLAYGROUND	0.517	0.0981	5.27	0.00	
O_VILLAGE	0.437	0.107	4.08	0.00	
O_WATER	0.304	0.0857	3.55	0.00	
O_WOODLAND	0.14	0.108	1.30	0.19	*
Number of estimated parameters:	34				
Number of observations:	995				
Null log-likelihood:	-2980.75				
Init log-likelihood:	-2980.75				
Final log-likelihood:	-2115.34				
Likelihood ratio test:	1730.825				
Rho-square:	0.29				
Adjusted rho-square:	0.279				
Diagnostic:	Convergence reached...				
Iterations:	8				
Significant. codes: ‘*’ > 0.05					

Table 4.25 *Model Results for Transit Trips, OSM Variables Only*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.038	0.00842	-4.51	0.00	
O_ALLOTMENT	-0.844	0.802	-1.05	0.29	*
O_AREA	0.6	0.255	2.35	0.02	
O_BEACHNCOAST	2.93	0.563	5.20	0.00	
O_CORRIDOR	-0.0112	0.389	-0.03	0.98	*
O_COUNTRYPARK	0.273	0.474	0.58	0.57	*
O_FARMLAND	-0.374	0.536	-0.70	0.49	*
O_IFGREEN	0.847	0.334	2.54	0.01	
O_MOUTAIN	0.713	0.998	0.71	0.47	*
O_PARKINCITY	0.792	0.319	2.48	0.01	
O_PLAYFIELD	0.054	0.43	0.13	0.90	*
O_PLAYGROUND	1.21	0.31	3.90	0.00	
O_VILLAGE	1.21	0.525	2.30	0.02	
O_WATER	0.484	0.373	1.30	0.19	*
O_WOODLAND	-1.01	0.648	-1.56	0.12	*
Number of estimated parameters:	34				
Number of observations:	84				
Null log-likelihood:	-251.642				
Init log-likelihood:	-251.642				
Final log-likelihood:	-126.278				
Likelihood ratio test:	250.728				
Rho-square:	0.498				
Adjusted rho-square:	0.363				
Diagnostic:	Convergence reached...				
Iterations:	13				
Significant. codes: ‘*’ > 0.05					

Table 4.26 *Model Results for Walking Trips, OSM Variables Only*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.0844	0.0049	-17.24	0.00	
O_ALLOTMENT	0.502	0.552	0.91	0.36	*
O_AREA	0.642	0.131	4.91	0.00	
O_BEACHNCOAST	2.6	0.419	6.20	0.00	
O_CORRIDOR	0.406	0.169	2.40	0.02	
O_COUNTRYPARK	0.7	0.225	3.10	0.00	
O_FARMLAND	0.45	0.346	1.30	0.19	*
O_IFGREEN	0.794	0.134	5.92	0.00	
O_MOUTAIN	1.21	0.463	2.62	0.01	
O_PARKINCITY	1.02	0.146	6.99	0.00	
O_PLAYFIELD	0.393	0.173	2.26	0.02	
O_PLAYGROUND	0.238	0.166	1.43	0.15	*
O_VILLAGE	0.12	0.292	0.41	0.68	*
O_WATER	0.59	0.216	2.73	0.01	
O_WOODLAND	-0.66	0.320	-2.06	0.04	
Number of estimated parameters:	34				
Number of observations:	1251				
Null log-likelihood:	-3747.66				
Init log-likelihood:	-3747.66				
Final log-likelihood:	-526.683				
Likelihood ratio test:	6441.956				
Rho-square:	0.859				
Adjusted rho-square:	0.85				
Diagnostic:	Convergence reached...				
Iterations:	13				
Significant. codes: '*' > 0.05					

As Table 4.24 shows, the driving model is converged after eliminating the GLUD variables, and the cycling model has been improved if comparing the Rho-suqare in Table 4.23 with the number in Table 4.19. In the next two sections, the model will be tested further through modifying alternatives by either distance or activity, the last two of the variables which have been reviewed in the beginning of this chapter.

4.6.2.2 Travel distance constrains

As an experiment to find out whether people would group choices by a range of distance when they make their choices, the number of available alternatives is narrowed down by setting a travel time limit for each transport mode. For example, in an observed trip, individual *n* has taken *X* minutes to travel to the selected destination; then, a time variable *Y* minutes is used as a limit. Alternatives are only considered to be available if travel to these destinations cost individual *n* longer than *X-Y* minutes and less than *X+Y* minutes. Travel mode used to calculate the travel time is same as the one selected by people *n* recorded in the MENE survey. Three different time limits are tested: 15 minutes, 30 minutes and 45 minutes. The results are shown in the following tables.

Table 4.27 *Summaries for Cycling Trips' Models by Different Travel Time Limit*

	All	15min	30min	45min
Number of estimated parameters:	41	41	41	41
Number of observations:	71	71	71	71
Null log-likelihood:	-212.697	-209.856	-212.697	-212.697
Init log-likelihood:	-212.697	-209.856	-212.697	-212.697
Final log-likelihood:	-35.707	-95.027	-153.712	-7.461
Likelihood ratio test:	353.98	229.658	117.971	410.472
Rho-square:	0.832	0.547	0.277	0.965
Adjusted rho-square:	0.639	0.352	0.085	0.772
Diagnostic:	Convergence reached...	Convergence reached...	Convergence reached...	Convergence reached...
Iterations:	20	14	14	25

.

Table 4.28 *Summaries for Driving Trips' Models by Different Travel Time Limit*

	All	15min	30min	45min
Number of estimated parameters:	34	41	41	41
Number of observations:	995	995	995	995
Null log-likelihood:	-2980.75	-2977.76	-2975.24	-2978.45
Init log-likelihood:	-2955	-2977.76	-2975.24	-2978.45
Final log-likelihood:	-2115.34	-803.314	-1111.24	-916.819
Likelihood ratio test:	1730.825	4348.888	3728	4123.264
Rho-square:	0.29	0.73	0.627	0.692
Adjusted rho-square:	0.279	0.716	0.613	0.678
Diagnostic:	Convergence reached...	Convergence reached...	Convergence reached...	Convergence reached...
Iterations:	45	26	13	34

For the cycling (Table 4.27) and driving (Table 4.28) models, judged by Rho-square, the model with the 45-minute limit gives the best Rho-square. In cycling models

(Table 4.27), the variable considered to be significant in the 45-minute model but not in the others is PARKINCIY. On the other hand, the variable MOUTAIN, which is suggested to be important in the non-limit model, is not anymore in the 45-minute limit model. Setting limits on driving models results not only in changing variables' significance level but also changing values if comparing the correlation coefficients in Table 4.32 with what has been recorded in Table 4.24: variables such as ALLOTMENT, BEACHNCOAST, COUNTRYPARK, IFGREEN and PARKINCITY changed their signs when estimated by different models.

Table 4.29 *Summaries for Transit Trips' Models by Different Travel Time Limit*

	All	15min	30min	45min
Number of estimated parameters:	41	41	41	41
Number of observations:	84	84	84	84
Null log-likelihood:	-251.642	-249.938	-251.642	-251.642
Init log-likelihood:	-251.642	-249.938	-251.642	-251.642
Final log-likelihood:	-110.086	-154.436	-147.155	-147.336
Likelihood ratio test:	283.11	191.003	208.974	208.61
Rho-square:	0.563	0.382	0.415	0.414
Adjusted rho-square:	0.4	0.218	0.252	0.252
Diagnostic:	Convergence reached...	Convergence reached...	Convergence reached...	Convergence reached...
Iterations:	20	11	11	11

Table 4.30 *Summaries for Walking Trips' Models by Different Travel Time Limit*

	All	15min	30min	45 mins
Number of estimated parameters:	41	41	41	41
Number of observations:	1251	1251	1251	1251
Null log-likelihood:	-3747.66	-2495.76	-3271.46	-3547.62
Init log-likelihood:	-3747.66	-2495.76	-3271.46	-3547.62
Final log-likelihood:	-498.593	-1341.75	-1844.06	-1609.73
Likelihood ratio test:	6498.136	2308.036	2854.809	3875.781
Rho-square:	0.867	0.462	0.436	0.546
Adjusted rho-square:	0.856	0.446	0.424	0.535
Diagnostic:	Convergence reached...	Convergence reached...	Convergence reached...	Convergence reached...
Iterations:	17	10	10	12

For transit and walking models, the 'non-limit' tests return the highest Rho-square as shown in Table 4.33 and Table 4.34. Among the tests for transit models, the correlation coefficients for some of the variables are inconsistent. For instance, in Table 4.33, ALLOTMENT, BEACHNCOAST, and FARMLAND changed their signs when the limits were added.

In conclusion, applying the time limits ‘15 minutes’ and ‘30 minutes’ did not improving any model results (Table 4.27 - Table 4.30). The cycling (Table 4.27) and driving (Table 4.28) models seem to be improved by using the ‘45 minutes’ travel time limits on the alternatives judged by Rho-square. However, the correlation coefficients of the variables mentioned in this section are inconsistent with the results in Section 4.3.2.6. Some results are neither rational nor compatible with previous studies. For example, if the correlation coefficient of ‘O_PARKINCITY’ in Table 4.32, under the ’45 min’ model, equals to -0.524, which means people who chose the drive to the destination less than 45 minutes’ way have found that parks in cities are a negative factor for their outdoor recreational trips in general. Given that the increases of Rho-square by adding travel time limits are scant, and the values of parameters are irreconcilable with previous outcomes, the model will be kept the same as was presented in Section 4.3.2.5 based on the tests done in this section. The effects of distance range according to people’s choice behaviours regarding outdoor recreation, which are not clear based on the available data at this stage.

Table 4.31 *Comparing Results of Cycling Trips by Different Travel Time Limit*

	All					15min					30min					45min				
Name	Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value	
B_TIME	-0.149	0.0277	-5.36	0.00		-0.385	0.0471	-8.19	0.00		-0.00434	0.00594	-0.73	0.46	*	-0.591	0.104	-5.71	0.00	
O_ALLOTMENT	-5.66	2.95	-1.92	0.05	*	-1.78	1.65	-1.08	0.28	*	-1.62	1.08	-1.50	0.13	*	-4.53	178	-0.03	0.98	*
O_BEACHNCOAST	-0.55	1.63	-0.34	0.74	*	0.739	0.729	1.01	0.31	*	0.154	0.638	0.24	0.81	*	-0.684	2.34	-0.29	0.77	*
O_CORRIDOR	1.77	0.678	2.62	0.01		1.9	0.425	4.47	0.00		1.67	0.344	4.87	0.00		3.07	1.43	2.15	0.03	
O_COUNTRYPARK	2.51	0.901	2.78	0.01		1.24	0.427	2.90	0.00		0.876	0.377	2.32	0.02		7.28	2.46	2.96	0.00	
O_FARMLAND	1.05	1.22	0.86	0.39	*	0.17	0.577	0.30	0.77	*	0.715	0.451	1.58	0.11	*	-0.775	2.03	-0.38	0.70	*
O_IFGREEN	0.658	0.584	1.13	0.26	*	-0.394	0.361	-1.09	0.27	*	-0.427	0.323	-1.32	0.19	*	-0.547	1.76	-0.31	0.76	*
O_MOUTAIN	5.88	1.57	3.74	0.00		4.74	1.01	4.70	0.00		2.1	0.67	3.14	0.00		6.96	5.89	1.18	0.24	*
O_PARKINCITY	-0.214	0.81	-0.26	0.79	*	-0.214	0.402	-0.53	0.60	*	0.0574	0.328	0.17	0.86	*	-4.49	2	-2.25	0.02	
O_PLAYFIELD	0.517	0.73	0.71	0.48	*	0.346	0.458	0.76	0.45	*	0.199	0.378	0.53	0.60	*	-1.57	2.5	-0.63	0.53	*
O_PLAYGROUND	-0.621	1.06	-0.58	0.56	*	0.128	0.503	0.25	0.80	*	1.04	0.415	2.50	0.01		2.07	3.18	0.65	0.51	*
O_VILLAGE	0.684	1.5	0.46	0.65	*	-0.0531	0.69	-0.08	0.94	*	0.797	0.562	1.42	0.16	*	3.3	2.77	1.20	0.23	*
O_WATER	1.35	0.994	1.36	0.17	*	0.85	0.455	1.87	0.06	*	0.684	0.372	1.84	0.07	*	4.25	2.19	1.94	0.05	*
O_WOODLAND	-0.811	1.21	-0.67	0.50	*	-0.111	0.637	-0.17	0.86	*	-0.216	0.515	-0.42	0.67	*	2.92	2.86	1.02	0.31	*

Table 4.32 *Comparing Results of Driving Trips by Different Travel Time Limit*

	All					15min					30min					45min				
Name	Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value	
B_TIME	-0.0526	0.00238	-22.12	0.00		-0.223	0.00668	-33.35	0.00		-0.395	0.0107	-36.79	0.00		-0.139	0.00451	-30.82	0.00	
O_ALLOTMENT	-0.391	0.218	-1.79	0.07	*	0.609	0.524	1.16	0.24	*	0.0594	0.382	0.16	0.88	*	1.58	0.644	2.46	0.01	
O_BEACHNCOAST	1.11	0.144	7.72	0.00		-1.01	0.238	-4.27	0.00		-0.435	0.194	-2.24	0.02		-1.5	0.226	-6.64	0.00	
O_CORRIDOR	-0.0646	0.0987	-0.65	0.51	*	-0.601	0.207	-2.90	0.00		-0.516	0.162	-3.18	0.00		-0.7	0.207	-3.38	0.00	
O_COUNTRYPARK	0.739	0.0964	7.67	0.00		-0.257	0.194	-1.32	0.19	*	0.307	0.144	2.13	0.03		-0.524	0.192	-2.73	0.01	
O_FARMLAND	0.106	0.111	0.96	0.34	*	0.116	0.227	0.51	0.61	*	0.333	0.160	2.08	0.04		0.46	0.239	1.92	0.05	*
O_IFGREEN	0.0559	0.0795	0.70	0.48	*	0.104	0.145	0.72	0.47	*	0.213	0.110	1.94	0.05	*	-0.0481	0.138	-0.35	0.73	*
O_MOUTAIN	0.828	0.141	5.86	0.00		1.15	0.268	4.29	0.00		0.806	0.210	3.83	0.00		1.21	0.270	4.47	0.00	
O_PARKINCITY	0.284	0.0894	3.18	0.00		-0.0792	0.158	-0.50	0.62	*	0.269	0.123	2.20	0.03		-0.0539	0.158	-0.34	0.73	*
O_PLAYFIELD	0.299	0.0925	3.23	0.00		0.312	0.180	1.73	0.08	*	0.507	0.137	3.71	0.00		0.0533	0.183	0.29	0.77	*
O_PLAYGROUND	0.576	0.105	5.51	0.00		0.833	0.207	4.01	0.00		0.657	0.166	3.95	0.00		0.803	0.195	4.11	0.00	
O_VILLAGE	0.551	0.115	4.80	0.00		1.31	0.238	5.48	0.00		0.527	0.164	3.22	0.00		1.83	0.256	7.14	0.00	
O_WATER	0.15	0.095	1.58	0.12	*	1.27	0.185	6.89	0.00		0.978	0.141	6.94	0.00		1.16	0.183	6.33	0.00	
O_WOODLAND	0.126	0.108	1.16	0.25	*	0.315	0.200	1.57	0.12	*	0.188	0.154	1.22	0.22	*	0.22	0.192	1.15	0.25	*

Table 4.33 *Comparing Results of Transit Trips by Different Travel Time Limit*

	All					15min					30min					45min				
Name	Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value	
B_TIME	-0.0368	0.00609	-6.03	0.00		0	1.80e+308	0.00	1.00	*	0	1.80e+308	0.00	1.00	*	0	81400000	0.00	1.00	*
O_ALLOTMENT	-1.11	0.998	-1.11	0.27	*	-2.08	0.791	-2.63	0.01		-1.75	0.789	-2.21	0.03		-1.45	0.825	-1.75	0.08	*
O_BEACHNCOAST	1.77	0.585	3.03	0.00		0.357	0.495	0.72	0.47	*	-0.0511	0.538	-0.10	0.92	*	-0.982	0.582	-1.69	0.09	*
O_CORRIDOR	0.589	0.412	1.43	0.15	*	0.853	0.377	2.26	0.02		1.39	0.405	3.45	0.00		1.31	0.424	3.10	0.00	
O_COUNTRYPARK	0.904	0.479	1.89	0.06	*	0.598	0.433	1.38	0.17	*	0.278	0.454	0.61	0.54	*	-0.0205	0.478	-0.04	0.97	*
O_FARMLAND	-0.27	0.668	-0.40	0.69	*	0.243	0.552	0.44	0.66	*	-0.172	0.54	-0.32	0.75	*	0.215	0.578	0.37	0.71	*
O_IFGREEN	0.38	0.345	1.10	0.27	*	0.237	0.285	0.83	0.41	*	0.169	0.309	0.55	0.58	*	0.0646	0.321	0.20	0.84	*
O_MOUTAIN	0.518	1.11	0.47	0.64	*	2.08	0.917	2.27	0.02		2.42	0.887	2.72	0.01		2.29	0.884	2.59	0.01	
O_PARKINCITY	0.656	0.384	1.71	0.09	*	0.953	0.326	2.93	0.00		1.19	0.325	3.65	0.00		1.02	0.351	2.91	0.00	
O_PLAYFIELD	0.411	0.451	0.91	0.36	*	0.481	0.387	1.24	0.21	*	0.492	0.392	1.26	0.21	*	0.615	0.397	1.55	0.12	*
O_PLAYGROUND	0.851	0.404	2.10	0.04		1.28	0.359	3.57	0.00		1.4	0.376	3.73	0.00		1.6	0.376	4.25	0.00	
O_VILLAGE	1.54	0.561	2.74	0.01		0.704	0.481	1.47	0.14	*	0.507	0.488	1.04	0.30	*	0.789	0.522	1.51	0.13	*
O_WATER	0.0166	0.417	0.04	0.97	*	0.567	0.349	1.63	0.10	*	1.01	0.351	2.87	0.00		1.09	0.361	3.03	0.00	
O_WOODLAND	-0.718	0.601	-1.20	0.23	*	-0.4	0.462	-0.87	0.39	*	-0.673	0.503	-1.34	0.18	*	-0.0831	0.511	-0.16	0.87	*

Table 4.34 *Comparing Results of Walking Trips by Different Travel Time Limit*

	All					15min					30min					45 mins				
Name	Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value		Value	Std err	t-test	p-value	
B_TIME	-0.0866	0.00394	-21.99	0.00		-0.111	0.00662	-16.83	0.00		-0.13	0.00419	-30.95	0.00		-0.147	0.00409	-35.85	0.00	
O_ALLOTMENT	0.0312	0.552	0.06	0.95	*	-0.461	0.296	-1.56	0.12	*	-0.718	0.256	-2.81	0.01		-0.845	0.302	-2.80	0.01	
O_BEACHNCOAST	2.35	0.45	5.22	0.00		0.66	0.228	2.90	0.00		0.448	0.194	2.31	0.02		0.51	0.202	2.53	0.01	
O_CORRIDOR	0.514	0.159	3.24	0.00		0.553	0.0994	5.56	0.00		0.523	0.0869	6.01	0.00		0.403	0.0919	4.39	0.00	
O_COUNTRYPARK	0.562	0.201	2.80	0.01		0.802	0.114	7.04	0.00		0.744	0.101	7.34	0.00		0.517	0.107	4.82	0.00	
O_FARMLAND	0.146	0.289	0.51	0.61	*	0.0403	0.169	0.24	0.81	*	0.0494	0.154	0.32	0.75	*	-0.103	0.156	-0.66	0.51	*
O_IFGREEN	0.58	0.14	4.15	0.00		0.5	0.0925	5.40	0.00		0.387	0.0809	4.78	0.00		0.447	0.0869	5.15	0.00	
O_MOUTAIN	1.28	0.371	3.46	0.00		0.666	0.263	2.53	0.01		0.848	0.215	3.95	0.00		0.886	0.228	3.89	0.00	
O_PARKINCITY	0.963	0.153	6.28	0.00		0.799	0.0987	8.09	0.00		0.899	0.0868	10.36	0.00		0.905	0.0916	9.88	0.00	
O_PLAYFIELD	0.223	0.173	1.29	0.20	*	0.249	0.105	2.37	0.02		0.252	0.0903	2.80	0.01		0.341	0.0965	3.53	0.00	
O_PLAYGROUND	0.209	0.185	1.13	0.26	*	0.528	0.112	4.73	0.00		0.662	0.0987	6.71	0.00		0.539	0.108	4.98	0.00	
O_VILLAGE	0.352	0.262	1.34	0.18	*	0.237	0.185	1.28	0.20	*	0.368	0.162	2.27	0.02		0.394	0.178	2.22	0.03	
O_WATER	0.785	0.216	3.64	0.00		0.334	0.129	2.60	0.01		0.389	0.109	3.56	0.00		0.283	0.117	2.42	0.02	
O_WOODLAND	-0.876	0.253	-3.46	0.00		0.0933	0.158	0.59	0.56	*	0.0248	0.136	0.18	0.86	*	0.226	0.144	1.57	0.12	*

4.6.2.3 Activities

In order to answer the question as to whether there is any correlation between the type of activities and demand of outdoor recreational trips, the dataset has been divided into groups by activity as shown in Table *4.9*, and then the trips distribution model was run on the new subsets. The results are shown in Table *4.35* - Table *4.38*.

Table 4.35 *Model Results for Anyone Walking without a Dog*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.0853	0.00499	-17.07	0.00	
O_ALLOTMENT	0.515	0.544	0.95	0.34	*
O_AREA	6.17E-07	1.28E-07	4.82	0.00	
O_BEACHNCOAST	2.62	0.385	6.82	0.00	
O_PATHS	0.414	0.169	2.45	0.01	
O_COUNTRYPARK	0.664	0.224	2.96	0.00	
O_FARMLAND	0.452	0.348	1.30	0.19	*
O_IFGREEN	0.8	0.134	5.99	0.00	
O_MOUTAIN	1.3	0.457	2.84	0.00	
O_PARKINCITY	1.02	0.147	6.93	0.00	
O_PLAYFIELD	0.364	0.172	2.12	0.03	
O_PLAYGROUND	0.217	0.165	1.32	0.19	*
O_VILLAGE	0.0806	0.297	0.27	0.79	*
O_WATER	0.686	0.214	3.20	0.00	
O_WOODLAND	-0.736	0.319	-2.30	0.02	
Number of estimated parameters:		34			
Number of observations:		1251			
Null log-likelihood:		-3747.66			
Init log-likelihood:		-3747.66			
Final log-likelihood:		-520.326			
Likelihood ratio test:		6454.671			
Rho-square:		0.861			
Adjusted rho-square:		0.852			
Diagnostic:		Radius of the trust region is too small			
Iterations:		82			
Significant. codes: ‘*’ > 0.05					

Table 4.36 *Model Results for Anyone Walking with a Dog*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.0852	0.0086	-9.91	0.00	
O_ALLOTMENT	-0.238	0.449	-0.53	0.60	*
O_AREA	5.06E-07	7.37E-08	6.86	0.00	
O_BEACHNCOAST	2.17	0.344	6.31	0.00	
O_PATHS	0.103	0.141	0.73	0.47	*
O_COUNTRYPARK	0.932	0.164	5.69	0.00	
O_FARMLAND	0.154	0.238	0.65	0.52	*
O_IFGREEN	0.581	0.128	4.55	0.00	
O_MOUTAIN	0.46	0.311	1.48	0.14	*
O_PARKINCITY	0.682	0.149	4.58	0.00	
O_PLAYFIELD	0.191	0.162	1.18	0.24	*
O_PLAYGROUND	-0.0143	0.18	-0.08	0.94	*
O_VILLAGE	0.339	0.262	1.29	0.20	*
O_WATER	0.664	0.176	3.76	0.00	
O_WOODLAND	0.316	0.252	1.25	0.21	*
Number of estimated parameters:		34			
Number of observations:		680			
Null log-likelihood:		-2037.1			
Init log-likelihood:		-2037.1			
Final log-likelihood:		-479.934			
Likelihood ratio test:		3114.328			
Rho-square:		0.764			
Adjusted rho-square:		0.748			
Diagnostic:		Radius of the trust region is too small			
Iterations:		69			
Significant. codes: ‘*’ > 0.05					

Table 4.37 *Model Results for Anyone Taking Part in Sports or Playing with Children, Including Any Kind of Formal and Informal Sports*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.0796	0.00689	-11.56	0.00	
O_ALLOTMENT	-0.479	0.285	-1.68	0.09	*
O_AREA	3.59E-07	6.80E-08	5.28	0.00	
O_BEACHNCOAST	2.01	0.283	7.13	0.00	
O_PATHS	0.0754	0.14	0.54	0.59	*
O_COUNTRYPARK	1.14	0.147	7.77	0.00	
O_FARMLAND	0.471	0.2	2.35	0.02	
O_IFGREEN	0.231	0.121	1.90	0.06	*
O_MOUTAIN	0.886	0.32	2.77	0.01	
O_PARKINCITY	0.51	0.127	4.01	0.00	
O_PLAYFIELD	0.726	0.124	5.85	0.00	
O_PLAYGROUND	0.955	0.132	7.22	0.00	
O_VILLAGE	0.462	0.196	2.36	0.02	
O_WATER	0.0173	0.155	0.11	0.91	*
O_WOODLAND	-0.419	0.188	-2.23	0.03	
Number of estimated parameters:		34			
Number of observations:		684			
Null log-likelihood:		-2049.08			
Init log-likelihood:		-2049.08			
Final log-likelihood:		-742.15			
Likelihood ratio test:		2613.862			
Rho-square:		0.638			
Adjusted rho-square:		0.621			
Diagnostic:		Radius of the trust region is too small			
Iterations:		70			
Significant. codes: ‘*’ > 0.05					

Table 4.38 *Model Results for Any Other Kinds of Activities*

Name	Value	Robust Std err	Robust t-test	p-value	
B_TIME	-0.0413	0.00347	-11.90	0.00	
O_ALLOTMENT	0.138	0.273	0.51	0.61	*
O_AREA	4.27E-07	3.78E-08	11.30	0.00	
O_BEACHNCOAST	2.04	0.171	11.88	0.00	
O_PATHS	-0.374	0.11	-3.41	0.00	
O_COUNTRYPARK	0.666	0.124	5.37	0.00	
O_FARMLAND	0.258	0.155	1.66	0.10	*
O_IFGREEN	0.493	0.0987	4.99	0.00	
O_MOUTAIN	0.695	0.212	3.28	0.00	
O_PARKINCITY	0.0341	0.102	0.33	0.74	*
O_PLAYFIELD	0.043	0.12	0.36	0.72	*
O_PLAYGROUND	0.594	0.125	4.77	0.00	
O_VILLAGE	0.48	0.14	3.42	0.00	
O_WATER	0.403	0.118	3.41	0.00	
O_WOODLAND	-0.115	0.164	-0.70	0.48	*
Number of estimated parameters:		34			
Number of observations:		646			
Null log-likelihood:		-1935.24			
Init log-likelihood:		-1935.24			
Final log-likelihood:		-1251.4			
Likelihood ratio test:		1367.696			
Rho-square:		0.353			
Adjusted rho-square:		0.336			
Diagnostic:		Radius of the trust region is too small			
Iterations:		65			
Significant. codes: '*' > 0.05					

The experiment is informational, because it shows the variations of preferred destinations for groups of people who chose different activities. For example, as shown in Table *4.35* the top three walking (without a dog) destinations are beach and coast, park in the city and mountain. For people who walked a dog (Table *4.36*), they tend to go to country parks more than mountains. If they have their children with them instead of a dog, as Table *4.37* shows, the playground replaced mountains, becoming the third-most popular destination after beach/cost and park.

Unfortunately, these models' results are not applicable, because none of above models is converged judged by the 'Diagnostic' (Table *4.35* - Table *4.38*). This means the value of coefficients cannot be trusted even they are rational. It will need more observations for all groups to make the activity models robust enough to be implemented. For this reason, the activity feature will not be added in the final model, but this could be a direction to work on in the future.

4.7 Final model

Based on the analysis above, the outdoor recreational trips are estimated following the structure as shown in Figure 4.21: the trip-generation function calculates the total number of trips from each origin, and then the modal choice function splits the total trips into four different groups by travel mode. Finally, the trip-distribution function locates the trips on each single destination. Modal choice and trips distribution are trained in the form of a multinomial logit model.

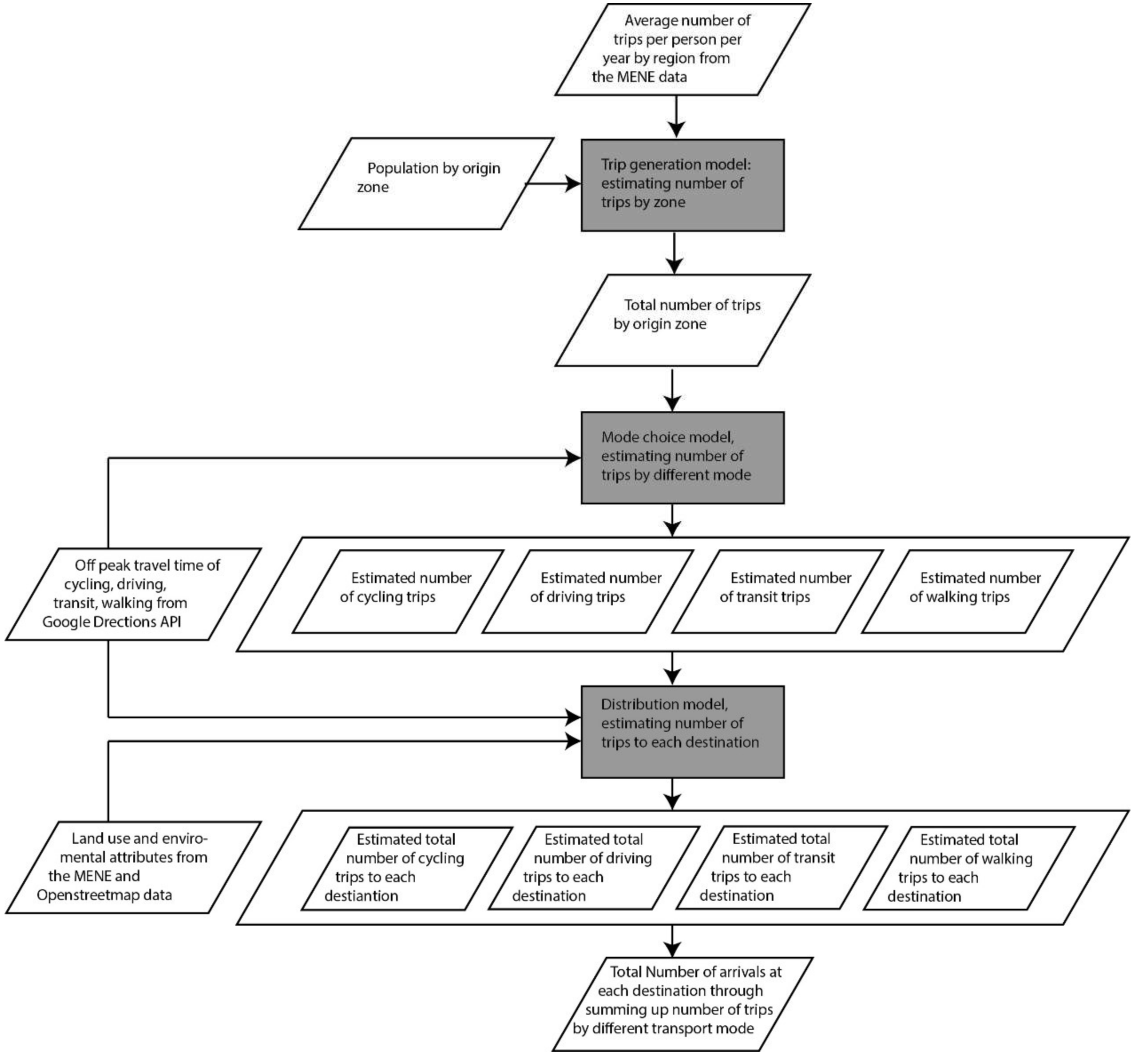


Figure 4.21. Schematic representation of modelling procedure.

4.7.1 Trip generation

The trip generation calculates the total number of trips generated from each origin, multiplying the mean of trips per person per year by the population of the corresponding zone. It can be written as:

$$T_j = T_i / Pop_i \times Pop_j \qquad \text{Equation 31}$$

Where T_j is total number of trips generated from origin zone *j*. T_i is the total number of trips in region *i* as recorded in the MENE data, Pop_i is the population of region *i*, Pop_j is population of neighbourhood *i*, and all the population numbers come from the 2011 census data.

4.7.2 Modal choice model

The first step estimates the total number of trips generated by each transport mode. This is achieved through a standard logit model taking a flowing format:

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_l e^{V_{nl}}} \qquad \text{Equation 32}$$

P_{ni} represents the probability of individual *n* choosing transport mode *i* to travel to the known destination observed by the MENE survey. The utility function is:

$$V_{ni} = a + \beta x_{ni} \qquad \text{Equation 33}$$

Where a is the constant, and β is the correlation coefficient to be estimated. X_{nk} denotes travel time from LOSA where individual *n* lives to the destination *k* through transport mode *i*. In our case, there are four different transportation modes: cycling, driving, transit and walking. Model results are shown in the following table:

Table 4.39 *Mode Split Model*

Name	Value	Robust Std err	Robust t-test	p-value
ASC_Car	0	fixed		
ASC_Cycling	-1.92	0.136	-14.19	0.00
ASC_Transit	-1.36	0.157	-8.66	0.00
ASC_Walking	1.74	0.175	9.96	0.00
B_TIME	-0.03	0.00496	-6.05	0.00
Number of estimated parameters:	4			
Number of observations:	2401			
Null log-likelihood:	-3445.722			
Init log-likelihood:	-3445.722			
Final log-likelihood:	-1815.574			
Likelihood ratio test:	3260.297			
Rho-square:	0.473			
Adjusted rho-square:	0.472			
Final gradient norm:	1.94E-03			
Diagnostic:	Convergence reached...			
Iterations:	8			
Significant. codes: '*' > 0.05				

Note: where ASC_Car is the constant for the car, the same applies to the other modes, constant for a can is fixed to be zero by default, B_TIME is the calibrated parameter for travel time.

4.7.3 Distribution model

When we know a total number of trips generated by each mode from origin *j*, the next stage is to estimate the number of trips to each individual site through the distribution model:

$$P_{nk} = \frac{e^{V_{nk}}}{\sum_m e^{V_{nm}}} \qquad \text{Equation 34}$$

P_{nk} represents the probability of individual *n* choosing destination *k* from *m* alternatives; the utility function is written as:

$$V_{nk} = \beta_{TRAVELTIME}TRAVELTIME_{nk} + \beta_{ALLOTMENT}ALLOTMENT_{nk} + \beta_{BEACHNCOAST}BEACHNCOAST_{nk} + \beta_{PATHS}PATHS_{nk} + \beta_{COUTRYPARK}COUTRYPARK_{nk} + \beta_{FARMLAND}FARMLAND_{nk} + \beta_{INGREEN}INGREEN_{nk} + \beta_{MOUTAIN}MOUTAIN_{nk} + \beta_{PARKINCITY}PARKINCITY_{nk} + \beta_{PLAYFIELD}PLAYFIELD_{nk} + \beta_{PLAYGROUND}PLAYGROUND_{nk} +$$

$\beta_{VILLAGE} VILLAGE_{nk} + \beta_{WATER} WATER_{nk} + \beta_{WOODLAND} WOODLAND_{nk} +$
$\beta_{AREA} AREA_{nk}$ Equation 35

There is no green spaces information in some of the LOSA zones based on either the MENE survey or OSM map. This does not necessarily equal no green space in the LOSA area. However, the green space (if there is any) could not be identified for outdoor recreational purposes. In other words, in some LOSAs, the area of green space does not equal zero, but the green space will not attract any outdoor recreational trips. Therefore, a size parameter S_k^{α} is introduced, and the final probability function can be written as:

$$P_{nk} = \frac{e^{V_{nk}} * S_k^{\alpha}}{\sum_m e^{V_{nm}} * S_j^{\alpha}}$$ Equation 36

Where S_k is the green space area of destination *k*, α is parameter which needs to be estimated practically. In this research, the α is set to be 0.4. Model results are shown below.

Table 4.40 *Distribution Model Results for Cycling Trips*

Name	Value	Std err	t-test	p-value	
O_ALLOTMENT	-4.22	2.23	-1.89	0.06	*
O_AREA	1.12	0.364	3.06	0.00	
O_BEACHNCOAST	1.34	1.55	0.86	0.39	*
O_PATHS	1.72	0.689	2.49	0.01	
O_COUNTRYPARK	2.85	0.878	3.25	0.00	
O_FARMLAND	0.274	1.1	0.25	0.80	*
O_IFGREEN	0.872	0.618	1.41	0.16	*
O_MOUTAIN	6.02	1.59	3.80	0.00	
O_PARKINCITY	-0.458	0.756	-0.61	0.54	*
O_PLAYFIELD	0.562	0.754	0.75	0.46	*
O_PLAYGROUND	-0.543	0.998	-0.54	0.59	*
B_TIME	-0.161	0.0307	-5.25	0.00	
O_VILLAGE	0.489	1.26	0.39	0.70	*
O_WATER	1.97	0.897	2.19	0.03	
O_WOODLAND	-1.31	1.21	-1.08	0.28	*
Number of estimated parameters:	34				
Number of observations:	71				
Null log-likelihood:	-212.697				
Init log-likelihood:	-212.697				
Final log-likelihood:	-35.319				
Likelihood ratio test:	354.755				
Rho-square:	0.834				
Adjusted rho-square:	0.674				
Diagnostic:	Convergence reached...				
Iterations:	15				
Significant. codes: ‘*’ > 0.05					

Table 4.41 *Distribution Model Results for Driving Trips*

Name	Value	Robust Std err	Robust t-test	p-value	
O_ALLOTMENT	-0.119	0.189	-0.63	0.53	*
O_AREA	0.358	0.0253	14.18	0.00	
O_BEACHNCOAST	1.73	0.138	12.54	0.00	
O_PATHS	-0.305	0.0868	-3.51	0.00	
O_COUNTRYPARK	0.973	0.0869	11.20	0.00	
O_FARMLAND	0.231	0.109	2.11	0.03	
O_IFGREEN	0.205	0.0775	2.64	0.01	
O_MOUTAIN	1.01	0.145	6.96	0.00	
O_PARKINCITY	0.137	0.0802	1.70	0.09	*
O_PLAYFIELD	0.265	0.0876	3.02	0.00	
O_PLAYGROUND	0.517	0.0981	5.27	0.00	
B_TIME	-0.0507	0.00487	-10.42	0.00	
O_VILLAGE	0.437	0.107	4.08	0.00	
O_WATER	0.304	0.0857	3.55	0.00	
O_WOODLAND	0.14	0.108	1.30	0.19	*
Number of estimated parameters:	34				
Number of observations:	995				
Null log-likelihood:	-2980.754				
Init log-likelihood:	-2980.754				
Final log-likelihood:	-2115.341				
Likelihood ratio test:	1730.825				
Rho-square:	0.29				
Adjusted rho-square:	0.279				
Diagnostic:	Convergence reached...				
Iterations:	8				
Significant. codes: ‘*’ > 0.05					

Table 4.42 *Distribution Model Results for Transit Trips*

Name	Value	Robust Std err	Robust t-test	p-value	
O_ALLOTMENT	-0.844	0.802	-1.05	0.29	*
O_AREA	0.6	0.255	2.35	0.02	
O_BEACHNCOAST	2.93	0.563	5.20	0.00	
O_PATHS	-0.0112	0.389	-0.03	0.98	*
O_COUNTRYPARK	0.273	0.474	0.58	0.57	*
O_FARMLAND	-0.374	0.536	-0.70	0.49	*
O_IFGREEN	0.847	0.334	2.54	0.01	
O_MOUTAIN	0.713	0.998	0.71	0.47	*
O_PARKINCITY	0.792	0.319	2.48	0.01	
O_PLAYFIELD	0.054	0.43	0.13	0.90	*
O_PLAYGROUND	1.21	0.31	3.90	0.00	
B_TIME	-0.038	0.00842	-4.51	0.00	
O_VILLAGE	1.21	0.525	2.30	0.02	
O_WATER	0.484	0.373	1.30	0.19	*
O_WOODLAND	-1.01	0.648	-1.56	0.12	*
Number of estimated parameters:	34				
Number of observations:	84				
Null log-likelihood:	-251.642				
Init log-likelihood:	-251.642				
Final log-likelihood:	-126.278				
Likelihood ratio test:	250.728				
Rho-square:	0.498				
Adjusted rho-square:	0.363				
Diagnostic:	Convergence reached...				
Iterations:	13				
Significant. codes: ‘*’ > 0.05					

Table 4.43 *Distribution Model Results for Walking Trips*

Name	Value	Robust Std err	Robust t-test	p-value	
O_ALLOTMENT	0.502	0.552	0.91	0.36	*
O_AREA	0.642	0.131	4.91	0.00	
O_BEACHNCOAST	2.6	0.419	6.20	0.00	
O_PATHS	0.406	0.169	2.40	0.02	
O_COUNTRYPARK	0.7	0.225	3.10	0.00	
O_FARMLAND	0.45	0.346	1.30	0.19	*
O_IFGREEN	0.794	0.134	5.92	0.00	
O_MOUTAIN	1.21	0.463	2.62	0.01	
O_PARKINCITY	1.02	0.146	6.99	0.00	
O_PLAYFIELD	0.393	0.173	2.26	0.02	
O_PLAYGROUND	0.238	0.166	1.43	0.15	*
B_TIME	-0.0844	0.0049	-17.24	0.00	
O_VILLAGE	0.12	0.292	0.41	0.68	*
O_WATER	0.59	0.216	2.73	0.01	
O_WOODLAND	-0.66	0.32	-2.06	0.04	
Number of estimated parameters:	34				
Number of observations:	1251				
Null log-likelihood:	-3747.661				
Init log-likelihood:	-3747.661				
Final log-likelihood:	-526.683				
Likelihood ratio test:	6441.956				
Rho-square:	0.859				
Adjusted rho-square:	0.85				
Diagnostic:	Convergence reached...				
Iterations:	13				
Significant. codes: '*' > 0.05					

At the final step, a total number of arrivals at each destination can be calculated as:

$$T_k = \sum_i \sum_j T_j \times P_{ji} \times P_{jk} \qquad \text{Equation 37}$$

T_k is the total number of trips to any site *k, i* is one of the four transport modes, *j* is the origin zone, T_j is total number of trips generated from the origin *j* by transport mode *i,* the probability of choose transport mode *i* is P_{ji} and the probability of selecting any destination *k* is P_{jk}. Per the values in Table 4.40–Table 4.43, travel time spent on trips is a definite drawback for all kinds of travel modes, but with different levels of significance. Cycling trips are the most affected by travel time (-0.161), followed by walking trips (-0.084). Driving (-0.05) and transit (-0.03) trips are less affected by travel time. As ASC values in Table 4.39 show, compared with driving (which the ASC value has been set to 0), walking is more preferred (1.71) as the travel method for outdoor recreation. Transit (-1.36) is less favoured than driving, and, again, cycling (-1.96) is the worst.

As correlation coefficients show in Table 4.40–Table 4.43, 'O_AREA' shows a significant positive value in all tables. It means the area of destination plays a decisive role in attracting all kinds of trips. It is apparently most important for those who choose to cycle (1.12) and least essential for people who decide to drive (0.358). This is rational because people who decide to cycle have had more freedom and require more space to move around inside the destination. On the other hand, for people who drive to outdoor recreational destinations, the characteristic of the destination is more important than area. Thus, they would spend more time at a relatively smaller place compared with cyclists. The difference of effects caused by area is small between transit and walking trips, with correlation coefficients of both at around 0.6. Beach and coast are found to be attractive most of the time. However, per the robust t-test in Table 4.40 for 'O_BEACHNCOAST', the correlation coefficient for cycling trips does not reach the statistically significant level. Country parks (O_COUNTRYPARK), mountains (O_MOUTAIN), and water features (O_WATER) all had a positive value in Table 4.40–Table 4.43. It means they all have been found attractive by everyone, although the effects on the transit traveller are insignificant (adjust p-values is smaller than 1.96 in Table 4.42). Informal green spaces (O_IFGREEN) express a very similar level of attractiveness to any other travellers (value equals to 0.8) except people who drive to the destinations (0.2).

Variables whose coefficients showing mixed results in Table 4.40– Table 4.43 include paths, cycling ways, and bridle ways (O_PATHS), and the results suggest that the paths, cycling ways, and bridle ways have a positive effect on attracting cycling (1.72) and walking (0.406) trips, but negative on driving trips as shown in Table 4.41, and unclear impacts on transit trips (Robust t-test is less than 1.96 in Table 4.42). Park in the city (O_PARKINCITY) is found attractive by those who choose to walk (1.02) and take public transport (0.79), but the opposite is found for cycling (-0.458). The value of 'O_PARKINCITY' for driving trips is positive (Table 4.41) but insignificant. Playground and sports field (including both formal and informal sports pitches) both play a decisive role in attracting driving trips (O_PLAYGROUND and O_PLAYFIELD in Table 4.41), but their effects on cycling trips are insignificant (Table 4.40). The playground is found to be attractive to individuals who take public transport and the same as sports field to walking trips. However, neither the effects of sports field to transit trips nor playground to walking trips has been found statistically significant (Table 4.42). The village (O_VILLAGE) seems only attractive to long trips through driving (0.437) or transit (1.21), particularly to people who choose to use public transport. However, the effects are not clear for short trips by cycling and walking (t-tests in Table 4.40 and Table 4.43 are less than 1.96). The only coefficient of woodland (O_WOODLAND) that reaches the significant level is the one for walking trips (Table 4.43), with a negative sign though. Farmland (O_FARMLAND) is the only variable found to be unimportant for all kinds of trips.

4.8 Summary

The focus of this chapter is calibrating the new travel demand forecasting model for outdoor recreational trips. The new model is built upon the raw data from the MENE survey. The case study area for model calibration consist of fourteen districts in the North-West region to facilitate the next two steps of this research: model validation and scenario tests. The variables studied in this research are informed by previous studies: travel time, land uses, land cover, population, the percentage of retired people and non-white ethnicity group, income, travel distance and activities.

The new model follows conventional transport modelling methods, formed by three parts: the trip-generation function, the modal choice and the trips distribution functions. The trip-generation function calculates the total number of trips generated from each origin, in this study the LSOA. This is done by multiplying the mean of outdoor recreational trips taken by individual lives in the North-West region per year, by the population of the LSOA. This number is then allocated to each destination using choice functions.

The choice functions for this research are calibrated through two steps. First, it defines the model structure. Three discrete choice model forms—multinomial logit model, nested logit model, and mixed logit mode—have been tested. As the results show, for outdoor recreational trips, travel time is as important as the characteristics of the green spaces, and people would normally choose their mode of transport first before deciding where to go. Therefore, the choice functions are first formed by modal choice function, a multinomial logit model. It divides all the trips generated from each origin zone into one of four transportation modes. Then, the trip-distribution function is also in the form of a multinomial logit model. The function estimates the number of trips to each destination based on travel time and the environmental characteristics of the destinations. The last part of calibration is finalising the exploratory variables; three groups of tests have been taken out, including removing insignificant variables, adding travel distance constrains and grouping data by activity. As the result shows, only the first process has clearly improved the model results.

Now the new outdoor recreation travel demand model has been trained, the next chapter will focus on validation, which is about proving how accurate the estimation is, and what the limitation is.

Chapter 5 Model Validation

5.1 Overview

Model validation is comparing the model results with the visit accounts from independent sources. This section is divided into two parts to answer two of the four research questions: is the estimation accurate enough and to what extent can the new model be transferred to destinations outside the case study area?

In the first part, the new model is applied on two nature reserves inside the model calibration area (Table *4.1*): Wigg Island and Wigan Flashes. The estimations from the new model are compared with the observations collected by the visitor counter in Wigg Island and the green space manager in the Wigan Flashes. In the second part, the application of the new model is extended to the area outside the model calibration zone, specifically, the ten English National Parks. The model results are then compared with the total number of visits to each of the ten National parks, which are reported in the final report of valuing England's national parks report (2013).

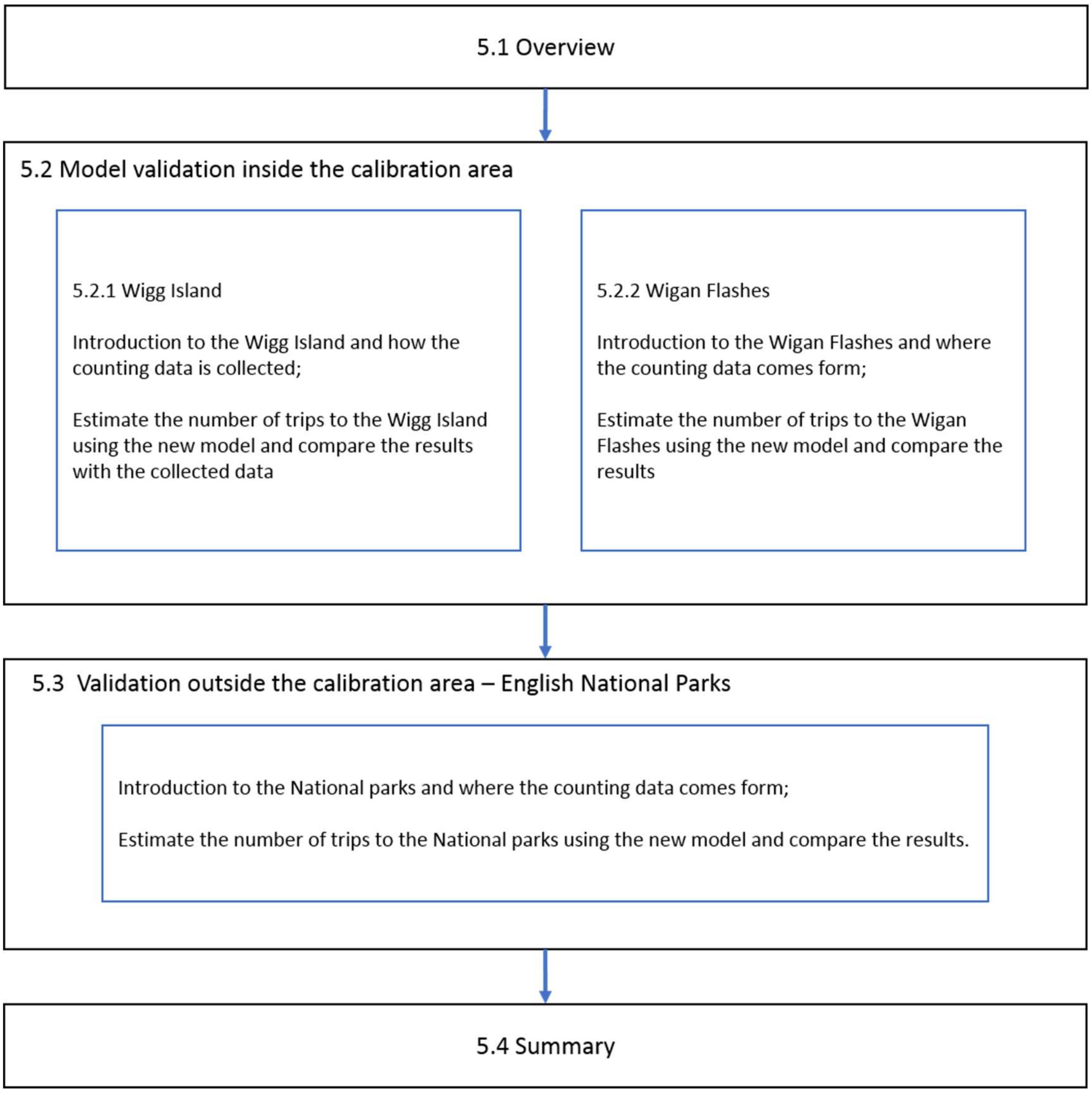


Figure 5.1. Structure of Chapter 5.

5.2 Model validation inside the model calibration area

The first group of tests was conducted on two of the nature reserves inside the model calibration area: the Wigg Island Nature Reserve and Wigan Flashes.

5.2.1 Wigg Island

Wigg Island is a local nature reserve near Runcorn in Halton, Cheshire, England. It lies between the River Mersey and the Manchester Ship Canal. It covers 23 hectares of reclaimed industrial land with wide footpaths and plenty of picnic tables. There are a number of bird hides and long views over the Mersey estuary and many wide-open spaces, great for woodland walks with trolls and sculptures. It is managed by the Halton Borough Council. It was made a Local Nature Reserve in 2004.

Two visitor counters were installed on the 16th of May 2012 at Wigg Island (Wigg Island People Counting System Monitoring Report, 2015). One of the two counters is to monitor the usage of the Visitor Centre; the other is to monitor the use of the nature reserve by people and cyclists. The counters register the number of activities that activate the sensors. The data are then calibrated to avoid double counting, for instance, to allow for staff working in the visitor centre and for people completing their walk using the same route. The latest data were for the year 2015; however, the Nature Reserve people counter had a nine-week gap in recording data, from 7th January to 14th April 2015 due to a damaged sensor. Therefore, the observation number of visits for 2014 is used for the validation purpose: 59,474 trips in total.

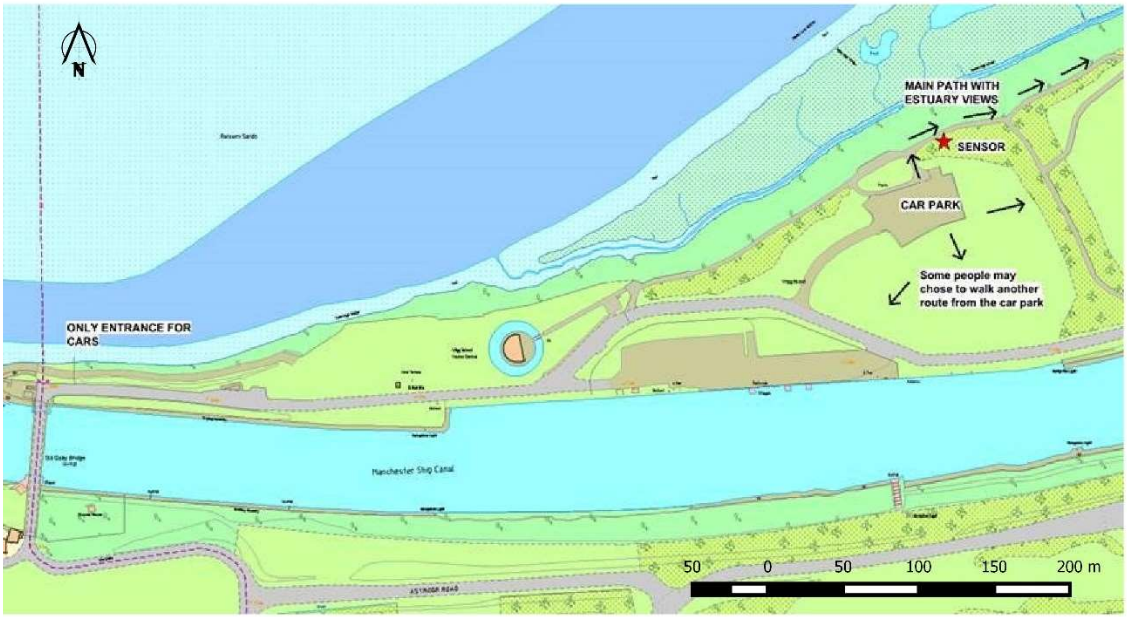


Figure 5.2. Location of the sensor, from the management team from the Mersey Gateway Crossings Board.

The boundary of origin is shown in Figure 5.3. The closer the origin is to the Wigg Island, the smaller neighbourhood it has applied to represent the boundary. Specifically, the Wigg Island is in Halton district and adjacent to the Warrington district. Therefore, when estimated the number of visits to the Wigg Island, the smallest neighbourhood boundary—LSOA—has been used as the origin in Halton and Warrington. For any other local authorities which are inside the model calibration area (Table *4.1*), the boundary of the local authority itself was used to represent the perimeter of the origin. The regional boundaries were used to represent anywhere outside the model calibration area in England. Figure 5.3 depicts the mean of trips per person per year from each origin area, and the result is showed in Table *5.1*.

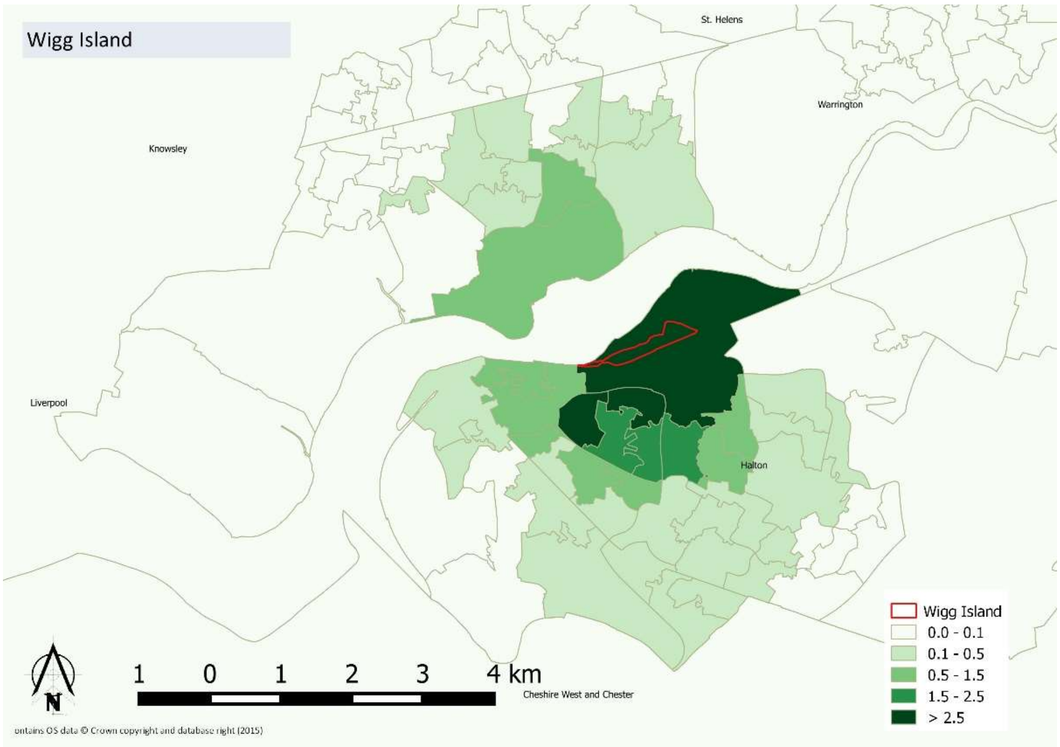


Figure 5.3. The mean of trips to Wigg Island per resident per year.

Table 5.1 *Simulation Results of Wigg Island.*

Wigg Island	Number of visits per year
Cycling	5,705
Driving	20,968
Transit	721
Walking	31,810
Total	59,204
Observed data	59,474
Residual	0
Difference (ratio)	0.5%

As Table *5.1* shows, this model gives a robust estimation on Wigg Island, with about a 0.5% difference when compared with the observed accounts. In the next section, one more estimate is conducted: the number of trips to another nature reserve inside the model calibration area, Wigan Flashes, is calculated using the new model.

5.2.2 Wigan Flashes

Wigan Flashes is a well-known wetland in the North-West region. It has been running for the last 14 years and managed by the Lancashire Wildlife Trust and Wigan Council. The site is 240 hectares in total, consisting of a group of eight shallow wetlands, formed initially as a result of mining subsidence, which extends south from near Wigan's town centre. Over time, the industrial landscape has evolved into a mixture of open water, reed bed, moss land and fenland[13]. The number of visits was observed by the site manager: the annual number of visits to Wigan Flashes in 2014 is around 100,000. However, this number is an estimation without any systematic data collection. Unfortunately, this is the best information we could have. In fact, detailed observation for any site is rarely existing and very expensive to obtain.

Figure 5.4. Snipped photo of Wigan Flashes site map, source: https://ttbirders.files.wordpress.com/2015/03/img_1314.jpg

The same method was applied to define the origins of the trips to Wigan Flashes. The difference is that is LOSA level zoning method is applied in Wigan only (Figure 5.5). The result is shown in Table *5.2*.

---

[13] http://www.wiganflashes.org/web/index.php?option=com_frontpage&Itemid=1

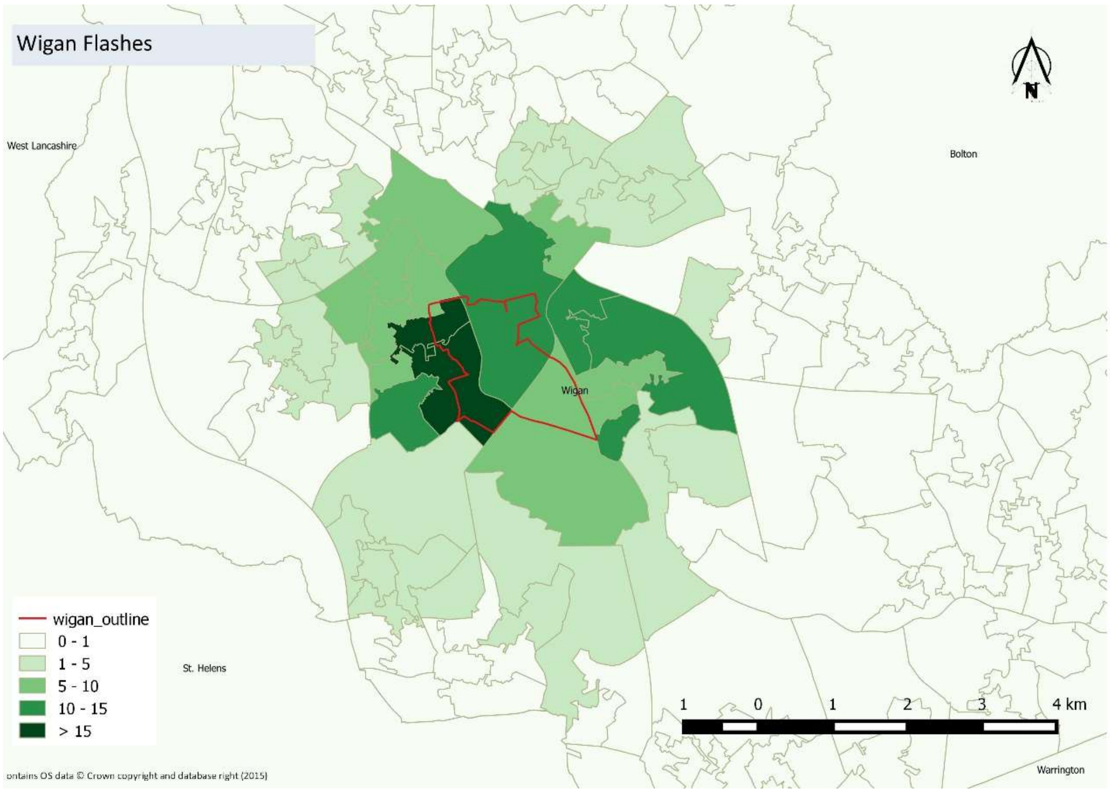


Figure 5.5 6. Mean of trips to Wigan Flashes per resident per year.

Table 5.2 *Simulation Results for Wigan Flashes*

Wigan Flashes	Number of visits per year
Cycling	7,367
Driving	19,356
Transit	436
Walking	96,776
Total	123,935
Observed data	100,000
Residual	0
Difference (ratio)	23.9%

Regarding the estimation for Wigan Flashes. There is a 23.9% difference comparing the number from the site manager. Given the number is also an estimation from the site manager. The only information we could learn from this test is, if Wigan Flashes is considered to be a standard outdoor recreational site in the North-West region, the total number of visits are attracted to the site per year, is 23,935 more than the manger's estimation, and this is suggested by the model. However, if we consider the manager's estimation is the figure true figure, the total number of visits is overestimated by the new model. It means the new model has not captured some features which are discouraging people from visiting Wigan Flashes. The standard practice in transport modelling to solve this kind of bias is to introduce a residual, which is placed at the end of the utility function, called an attraction residual (Hollander, 2016). In this case, if an attraction residual is equalling -0.22 is added, the difference is reduced to 0.7% (Table *5.3*).

Table 5.3 *Simulation Results for Wigan Flashes with an Attraction Residual*

Wigan Flashes	Number of visits per year
Cycling	5,967
Driving	15,534
Transit	350
Walking	78,894
Total	100,744
Observed data	100,000
Residual	-0.22
Difference (ratio)	0.7%

To answer the second research question regarding the accuracy of the estimation, two tests have been approved. The new model can make robust estimations for the outdoor recreational sites, although it might need some adjustments through attraction residuals. So far, the tests are staying inside the model calibration area. In the next section, the tests will be carried out on bigger green spaces outside the calibration area—the ten English National Parks. Each of them covers a few local authorities, and some of them even cross regions.

5.3 Validation of the sites are outside the calibration area – the English National Parks

The next group of tests explore to what extent the new model can be transferred to destinations outside the case study area. In this section, ten English National Parks were used to validate the new model. The National Parks are valued as national assets; they cover more than 1.2 million hectares (9.3% of the total land area) and comprise some of the highest quality landscapes and wildlife habitats (National Parks England 2013). National Parks form important centres for tourism and recreation and education. This is reflected in the large numbers of visitors they receive.

England's National Park Authorities have commissioned studies which have used the Scarborough Tourism Economic Activity Monitor (STEAM) to estimate the volume of visits.

Figure 5.7. Locations of National Parks (National Parks England 2013).

The Scarborough Tourism Economic Activity Monitor is derived from a spreadsheet model developed by David James and Frank Hart in the process of developing a ten-year tourism policy for the province of Saskatchewan, Canada, in 1981. The model was first to run on behalf of the Scarborough Borough Council in 1990. In 1991, North Yorkshire County Council, together with the District Councils in the county, embarked on a pilot programme to evaluate the now-named 'Scarborough/Scottish Tourism Economic Activity Monitor'. At the same time, STEAM was adopted by many local authorities in England, Scotland and Wales[14]. STEAM's estimations on annual visits to England's National Parks are shown in Table 4.44.

Table 5.4 *Estimated Numbers of Visits to England National Parks and Estimated Number of Visits*

National Park	Year of Estimate	Day trips	Staying visits	Total visits
The Broads	2011	6,308,000	602,800	6,910,800
Dartmoor	2011	2,042,000	234,000	2,276,000
Exmoor	2011	1,060,700	266,500	1,327,200
Lake District	2010	12,960,630	2,263,200	15,223,830
New Forest	2009	3,161,000	360,000	3,521,000
North York Moors	2011	5,099,650	477,920	5,577,570
Northumberland	2011	1,290,200	84,490	1,374,690
Peak District	2011	7,954,000	733,000	8,687,000
South Downs	2011/12	44,316,000	1,992,000	46,308,000
Yorkshire Dales	2011	3,117,000	410,000	3,527,000
Total		87,309,180	7,423,910	94,733,090

Note: Source: Valuing England's National Parks (National Parks England 2013)

Since the National Parks cover a relatively large area compared with the local natural reserves studied above, the Middle Super Output Area (MSOA) zoning system was used at the places where inside the boundaries of National Parks instead of LSOA. For any area outside the boundary of the national park but in the same region as where the national park sits, perimeters of local authorities were used to define the origins. The rest of England is modelled at the regional level.

The new model is built upon observations, whereby the majority travelled less than two-and-a-half hours (Figure 4.5). Therefore, the modelled results are expected to match the numbers under the 'Day Trips' column in Table *5.4*. Modelled results without attraction residuals are shown in Table *5.5*.

As Table *5.5* shows, the estimation is more accurate for the parks where are closer to the calibration area (i.e., Peak District, Yorkshire Dales, Northumberland, North York Moors), and the difference of modelled results compared with the STEAM results are less than 10%. The exception is the Lake District, where the new model clearly underestimated the attractiveness of the Lake District, given that the Lake District is one of the most famous outdoor recreational destinations in England, and the

[14] http://mediafiles.thedms.co.uk/Publication/LM/cms/pdf/STEAM%20OVERVIEW~%20Eng-Wal-NI.pdf

distinction of the Lake District is the kind of feature that cannot be easily captured by the model which is built upon the average basis. However, as shown in Table *5.6*, adding the attraction residuals, the absolute value of which are all less than 3, can make the model recalculate the total trips with less than 0.2% difference.

On the other hand, the model makes less accurate estimations for the parks further away from the model calibration area as shown in Table *5.5* (i.e., The Broads, Dartmoor, Exmoor, New Forest, South Downs). One emerging limitation is that the new model is built upon the data collected from individuals who lived in the North-West region as described in Chapter 4. The behaviours of people who live in the other regions might be significantly different. In particular in these regions which are in the south of England. However, the bias can be reduced in one of following methods. First, given that the MENE is a nation-wide survey, the database used to calibrate this model can be easily extended to the national level if required. Alternatively, a separate mode can be trained using the same method as presented in this research, albeit with the baseline data selected from people who live in the desired region. Due to the time and budget limits, the method is applied in this study by introducing the attraction residuals as for Wigan Flashes. The results are shown in Table *5.6*.

Table 5.5 *Estimated Number of Trips to National Parks per Year*

National Park	Cycling	Driving	Transit	Walking	Modelled all modes	Reported Total	Difference (Million)	Ratio (%)
The Broads	151,381	1,863,687	135,374	3,267,444	5,417,886	6,308,000	-0.89	14
Dartmoor	184,91	1,549,036	26,963	866,851	2,461,341	2,052,000	0.41	20
Exmoor	8,399	1,271,069	14,441	351,354	1,645,264	1,060,700	0.58	55
Lake District	114,027	4,122,801	122,738	3,174,067	7,533,633	12,960,630	-5.43	42
New Forest	69,649	1,976,648	112,908	1,615,912	3,775,117	3,161,000	0.61	19
North York Moors	112,127	1,723,576	164,126	3,216,395	5,216,225	5,099,650	0.12	2
Northumberland	8,324	477,277	4,429	787,927	1,277,957	1,290,200	-0.01	1
Peak District	256,352	4,935,201	212,598	3,319,431	8,723,582	7,950,000	0.77	10
South Downs	765,317	25,600,699	716,058	9,654,425	36,736,500	44,316,000	-7.58	17
Yorkshire Dales	88,735	853,617	64,879	1,862,422	2,869,653	3,117,000	-0.25	8

Table 5.6 *Estimated Number of Trips to National Parks, with Attraction Residuals*

National Park	Cycling	Driving	Transit	Walking	Modelled total	Residuals	Reported Total	Difference (Million)	Ratio (%)
The Broads	169,830	2,565,159	182,119	338,9420	6,306,528	0.52	6,308,000	-0.001	-0.02
Dartmoor	15,497	1,150,419	15,844	865,736	2,047,495	-0.65	2,052,000	-0.005	-0.22
Exmoor	6,851	695,959	6,774	351,330	1,060,914	-0.85	1,060,700	0.000	0.02
Lake District	132,401	9,287,210	199,783	3,318,289	12,937,683	1.46	12,960,630	-0.023	-0.18
New Forest	60,029	1,406,289	80,341	1,611,933	3,158,591	-0.45	3,161,000	-0.002	-0.08
North York Moors	110,986	1,624,771	156,200	3,208,606	5,100,563	-0.08	5,099,650	0.001	0.02
Northumberland	8,325	490,578	4,503	787,927	1,291,332	0.03	1,290,200	0.001	0.09
Peak District	237,239	4,279,585	186,159	3,246,984	7,949,967	-0.17	7,954,000	-0.004	-0.05
South Downs	832,841	32,519,190	925,218	10,023,011	44,300,260	0.44	44,316,000	-0.016	-0.04
Yorkshire Dales	92,283	1,078,900	77,758	1,863,168	3,112,109	0.24	3,117,000	-0.005	-0.16

5.4 Summary

In this chapter, to answer the second and third research questions, the new model has been applied to estimate the number of trips to two groups of outdoor recreational sites. The first group of the destinations is in the calibration area—Wigg Island and Wigan Flashes. The estimations made by the model were compared with the observed accounts from the management teams. As the results show, the new model makes decent estimations on Wigg Island, with a 0.5% difference compared with the visiting accounts reported by the people-counting monitors; on Wigan Flashes, the difference between the estimation and the number reported by the manager was 23.9%. By adding an attraction residual (-0.22) on the utility function, the difference was reduced to 0.7%.

In the second group of tests, the new model was implemented on the ten English National Parks, in order to explore to what extent the new model can be transferred to destinations outside the model calibration area. The model shows more robust estimations for those parks inside or close to North-West region, except the Lake District, where the visiting account is hugely underestimated. This means the Lake District is unique; hence, the standard practice to solve this kind of problem in transport modelling is to add an attraction residual to the utility function. On the other hand, results for those National Parks further away from the North-West region are less accurate. This is partly because the new model has been trained on baseline data selected from individuals who live in the North-West region; people who live in other regions might behave differently. Although there is a number of methods to deal with this kind of bias, due to time and budget limits, this research followed the fast solution mentioned above—adding an attraction residual. When the utility was adjusted using attraction residuals, the difference between modeled results and numbers from the National Parks report could be reduced to less than 0.3%.

The tests in this chapter indicate that this new model has great potential for estimating the travel demand for outdoor recreational trips. It offers robust estimations on the destinations inside or close to the model calibration area. Also, this model can be implemented on any site within England, although some adjustments will be needed. Except for the use of attraction residuals, the model results can also be improved by either expanding the baseline data to the national level or building a separate model for the desired region using the same method applied for calibrating this model. The MENE survey data have covered the whole England area, and this method has been approved to work successfully in the North-West region; there is no reason why this technique cannot be transferred to any other region using data from the open-source data used in this research.

Appearently, the model has a limitation regarding behaviours of people from all regions. However, it does not overthrow the fact that this new model has shown great potential to forecast the travel demand for any outdoor recreation destination. And it is particularly robust in the area where the data used for calibration were collected. Therefore, it is worth in the next chapter to discuss how this new model can be applied to assist planning and design activities. In Chapter 6, the new model will be applied to test a series of scenarios. These scenarios are developed for the Upper Mersey Estuary (UME), the model calibration area.

Chapter 6 Scenario Tests

6.1 Overview

This chapter aims to use the model developed in the last section to assess the implications of alternative patterns of future land developments in the Upper Mersey Estuary (UME) area. A series of tests were conducted, which led to three scenarios developed by Dr Andrea Drewitt from the Ecosystem and Environment Research Centre at the University of Salford. First, a trend scenario is defined (Business As Usual), which represents a continuation of the status quo. This was then compared with two more extreme situations. One focuses on booming economic development (Development Boom), by increasing the availability of land for commercial development and residential land use in and around the UME. A third one aims to better protect the environment through limiting commercial and residential development and promoting investments of green infrastructure (i.e., Nature is the Key).

In applying the new model, this study will estimate the changes of travel demand to the three of the outdoor recreational destinations on the UME site: two existing nature reserves—Wigg Island and Moore Nature Reserve—and a new proposed park built on an existing landfill site: Arpley Country Park.. These changes result from the individual planning and design interventions. The above three scenarios are compared through modelling the number of trips by transport mode to three destinations.

In section 6.2, the scenario design and assumptions of inputs are presented. In order to use the model to make assessments, these scenarios have to be interpreted in the form of changes of variables. This includes reducing the size of green spaces, adding new alternatives, varying environmental characteristics and adjusting the residential population in the catchments of the green spaces. All results are presented in section 6.3. Finally, in section 6.4, the main findings from the model applications are summarised. The process of model applications is shown in Figure 6.1.

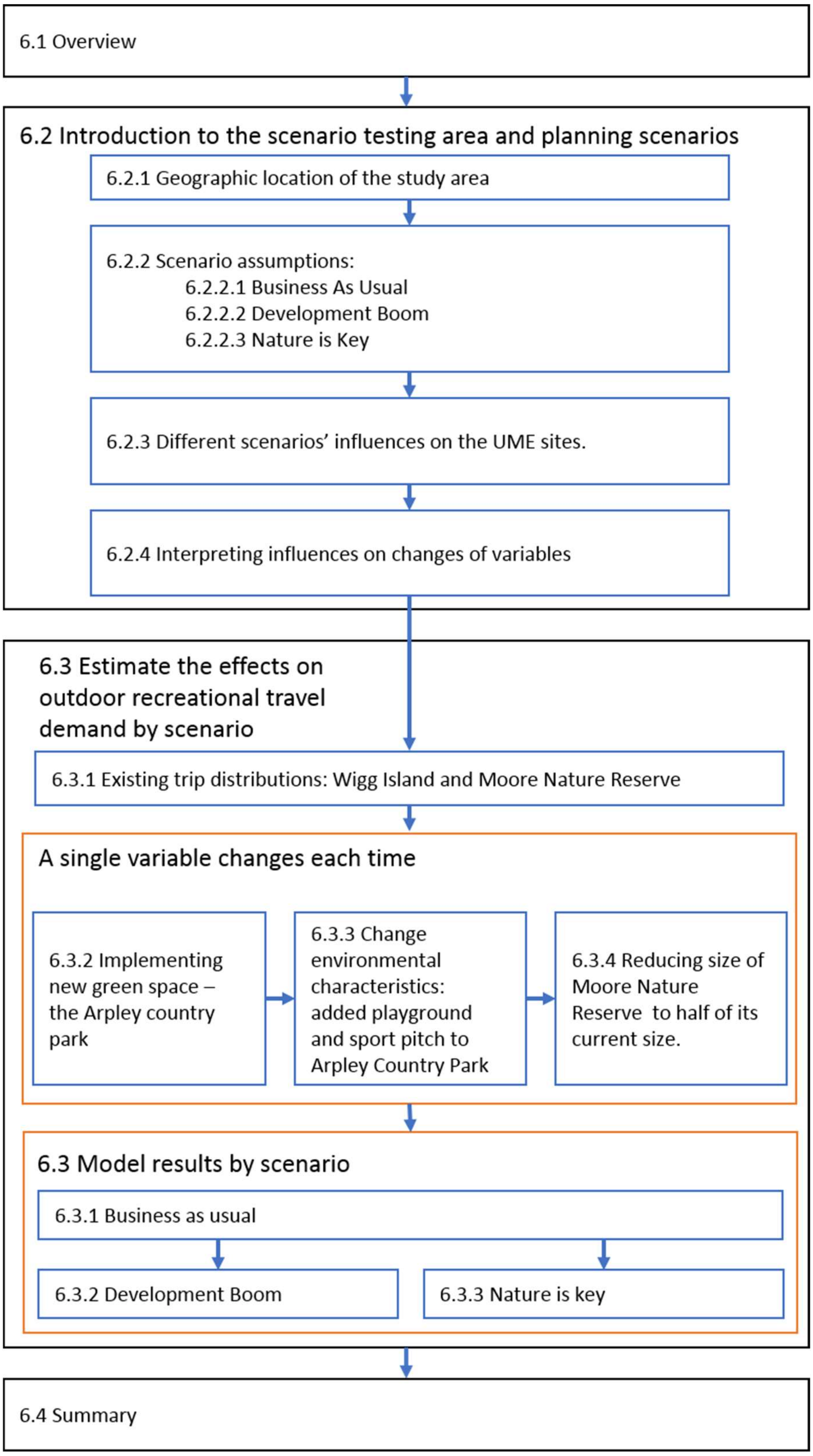


Figure 6.1. Structure of Chapter 6.

6.2 Scenarios

6.2.1 The Upper Mersey Estuary (UME) area.

The Mersey Estuary area starts from the upper tidal limit of Howley Weir in Warrington to the sea and stretches for a distance about 50 kilometres. It can be divided into four distinct zones: the Upper estuary, the Inner estuary, the Narrows and the Outer estuary. The Upper estuary is between Warrington and Runcorn; it is narrow and consists mostly of a single, meandering channel. Figure 6.2 shows the case study area and its components where the scenarios will be implemented.

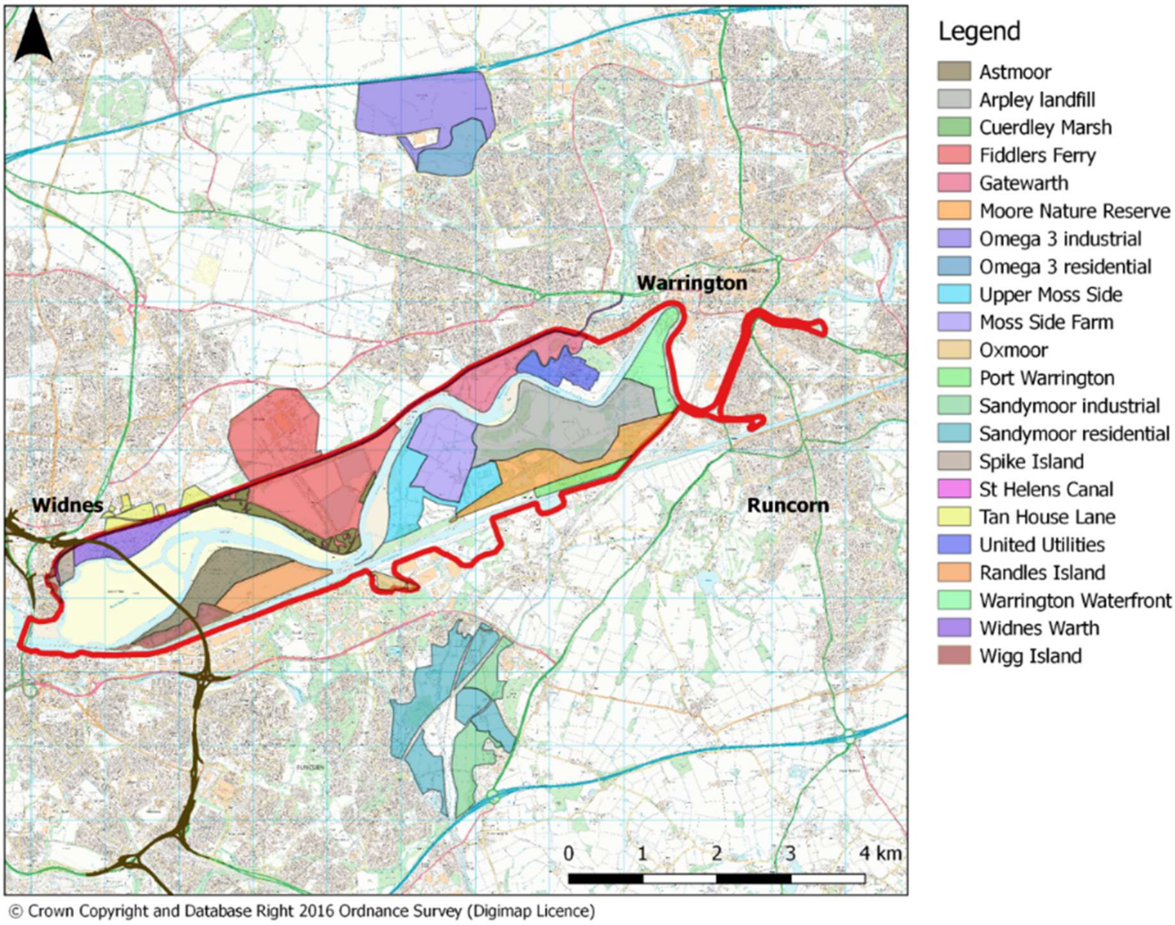


Figure 6.2. Upper Mersey Estuary site components. Source: Drewitt (2016).

6.2.2 Scenario assumptions

The three scenarios are Business as Usual 2044, Development Boom 2044 and Nature is the Key 2044. These were developed by Drewitt for a doctoral study in the Ecosystem and Environment Research Centre at the University of Salford. The scenarios cover the remaining years of the 30-year planning period, beginning with the construction of the Mersey Gateway in 2014. The scenarios envision different alternative futures for the UME in Halton by 2044 arising from changes in land use and environmental characteristics.

These scenarios have been identified through several means. First, policy reviews of the borough councils planning documents and core strategies give an impression of how the borough is planning to develop. Moreover, work with the stakeholders in the UME over the past years offers a thorough understanding of the dynamics by which the estuary was formed. As part of the scenario formation, several assumptions have to be formulated. All scenarios are based on these assumptions (Drewitt, 2016).

Firstly, population growth is average; i.e., it increases slightly but without significant migration in to or out of the UME. The strategy documents from Halton and Warrington support this assumption. Secondly, many (brownfield) sites in the boroughs are contaminated development sites and are therefore restricted development sites, making them less attractive to investors than sites without contamination issues. In addition, political factors that are of national importance and which could influence national and local policies (i.e., Brexit) cannot be considered within the scenarios.

6.2.2.1 Business as Usual 2044 (BAU)

This scenario explores how the estuary could look if the status quo is maintained. Under this situation, land use plans remain in place, and contemporary trends are still viable in 2044. The provision of housing is a high priority to the boroughs. By 2044, most of the housing and commercial development that had been proposed in the local authorities' core strategies has been delivered. New developments will be subjected to Green Belt and other binding legal obligations. The connection between the big cities of Liverpool and Manchester is taking place, but the idea that smaller places like Widnes, Runcorn and Warrington can thrive through local infrastructure remains. Science and technology advance at a moderate speed, offering more possibilities for developers and building sites. Some changes are made to the Green Belt in and around the UME. All nature designations have been kept and Arpley landfill site has been developed as per plan into a recreational nature site. Moore Nature Reserve is kept as the designated Local Nature Reserve.

6.2.2.2 Development Boom 2044 (DB)

In this scenario, economic development is the central aspect of the planning period. Economic growth is driven by a consumption-oriented society which devotes a lot of resources to the production of goods. The central position of the UME in the North-West is recognised as a hub for regional and national development importance. Changes in land-use regulations (i.e., Green Belt) increase the availability of land for commercial development and residential land use in and around the UME. Through the advancement in technology, brownfield sites might become available for development through improved techniques dealing with contaminated soils. The opening of the Green Belt for further development will decrease natural corridors, and there is no significant focus on nature protection and conservation. The Mersey Gateway is an essential infrastructural link and will attract development to the area.

6.2.2.3 Nature is the Key 2044 (NK)

The natural enviroment is the focus point in local and national development. The laws develop in a direction whereby environmental protection is a crucial issue, and a focus on conservation and protection is an essential part of the legislation. Access to green spaces plays a role in the development of new housing. People enjoy nature and appreciate green spaces. The use of local green spaces has become a valuable resource for society to spend free time. Outdoor activities and environmental education are encouraged. No new land is being used for the development of new industrial sites. Brownfield and other previously developed sites are used for the necessary site development. The investment in green infrastructure has been promoted. This includes easier access to funding and long-term planning opportunities, whereby transport modes become more sustainable, and travel distances are optimised. The Mersey Gateway is working within its capacity. Sustainable use of resources benefits from the advancements in technology. The opinion of local people is essential in decision making. More detail on this scenario can be found in Appendix C.

6.2.3 Interpreting the influences on changes in variables

The new model is used to estimate the number of trips to an outdoor recreational site on the basis of land use and the demographic changes. The sites to be influenced, as proposed in the scenarios, are Arpley, Moore Nature Reserve, St Helens Canal, Tan House Lane, Warington waterfront, Port Warrington Sandymoor and Omega 3. Before the new model can be used to test the scenarios, the influences need to be interpreted in the form of the changes of variables. Firstly, in the following tests, wherever new developments are concerned, the population changes are estimated based on the assumption that the mean of house density in this area are 50 dwellings per hectare, with 2.5 people in each dwelling. When it says partially developed, it was considered that only 50% of the total area is developed, and small development equals to 25%. When it says mix-used development, it means 25% of the area is used for residential development. The detailed changes in variables applied in the following test are listed in Table *6.1*.

Table 6.1 *Variables Changes*

Site	Area (km^2)	Business as Usual		Development Boom		Nature is Key	
		Changes	Variable	Changes	Variable	Changes	Variable
Arpley	1.65	landfill to country park	Path =1, Water=1, Countrypark =1, Parkincity =1	landfill to country park	Path =1, Water=1, Countrypark =1, Parkincity =1	landfill to coutry park	Path =1, Water=1, Countrypark =1, Parkincity =1
Moor Nature Reserve	0.88	Parts will be taken up for development	None	Stay the same	Area decreased: 0.88 km2 to 0.44 km2	Stay the same	
St Helens Canal	0.17	Stay the same		Stay the same	None	natural reserve	Path =1, Water=1, Countrypark =1, Parkincity =1
Tan Houose Lane	0.27	Partial development for mixed use	Population increased: 406	Full development of mixed use	Population increased: 813	no development	
Warington waterfront	0.44	residential housing	Population increased: 5,500	residential housing	Population increased: 5,500	no development	
Por Warrington	0.19	full development	Population increased: 232	full development	Population increased: 232	no development	
Sandymoor	1.49	residential housing	Population increased: 18,500	residential housing	Population increased: 18,500	residential housing	
Omega 3	0.43	residential housing	Population increased: 5,250	residential housing	Population increased: 5,250	residential housing	

6.3 Applications of the model

This section covers the model applications on Wigg Island, Moore Nature Reserve, and Arpley Country Park. The application steps are shown in Figure 6.3. Firstly, only one attribute is changed at a time. This is to estimate the change of travel demand caused by adjusting a single variable. In section 6.3.1, the current number of trips to Wigg Island and Moore Nature Reserve is estimated and mapped first. In 6.3.2, a new country park, Arpley Country Park is implemented as proposed by all the scenarios. The number of trips to Arpley Country Park is then estimated based on the assumption that Arpley Country Park will have the same characteristics as Wigg Island and Moore Nature Reserve do. In 6.3.3, two features are added to Arpley Country Park—a playground and a playing field. Then, the study examines the number of visits due to the instalment of these two new facilities. In section 6.3.4, the area of Moore Nature Reserve is reduced to half of its size as proposed by the Development Boom scenario. This is also the last part of the single variable changes, and, finally, in 6.3.5, the population changes caused by proposed new developments are incorporated into the estimations, and the final estimations for each scenario are presented by tranpsort mode.

The model is used to estimate the total number of trips to selected destinations from the places all over England. There will be too many origins if Lower Super Output Areas (LSOAs) were used to represent origins in England. Also, LSOA is too small a unit to study areas too far away from the destinations, because the travel effort required to go to the selected destination will outweigh the attractions. It means enviromental characteristics do not matter when travel time is longer than a certain limit, and these limits are different among travel modes. Therefore, the LSOAs are only used as the origins for the districts where the UME area sits (i.e., Warrington, Halton). For any places outside these two districts but inside the model calibaration area (Figure 4.x), the local authority boudary is used as the outline of the origin. The rest of England is divided by region. For both districts and regions, their gematric centroids are used to represent the starting point of the individual’s outdoor recreational trip.

One problem arising from this method of defining origins is that the bigger origin is, the larger number of trips will be generated regardless of distance sometimes. This is simply because the bigger origin has a larger population. For example, in comparing a LSOA of

Halton with the rest of the Manchester district, although the trip density (the mean of trips per person per year) will be higher in the LSOA closer to the UME, in terms of the total number of trips to the destination, the Manchester district will always win. However, this is not telling the full story of the trip distributions. Therefore, the model results are mapped by trip density (the mean of trips per person per year) instead of the total number of trips from each origin as the model outputs.

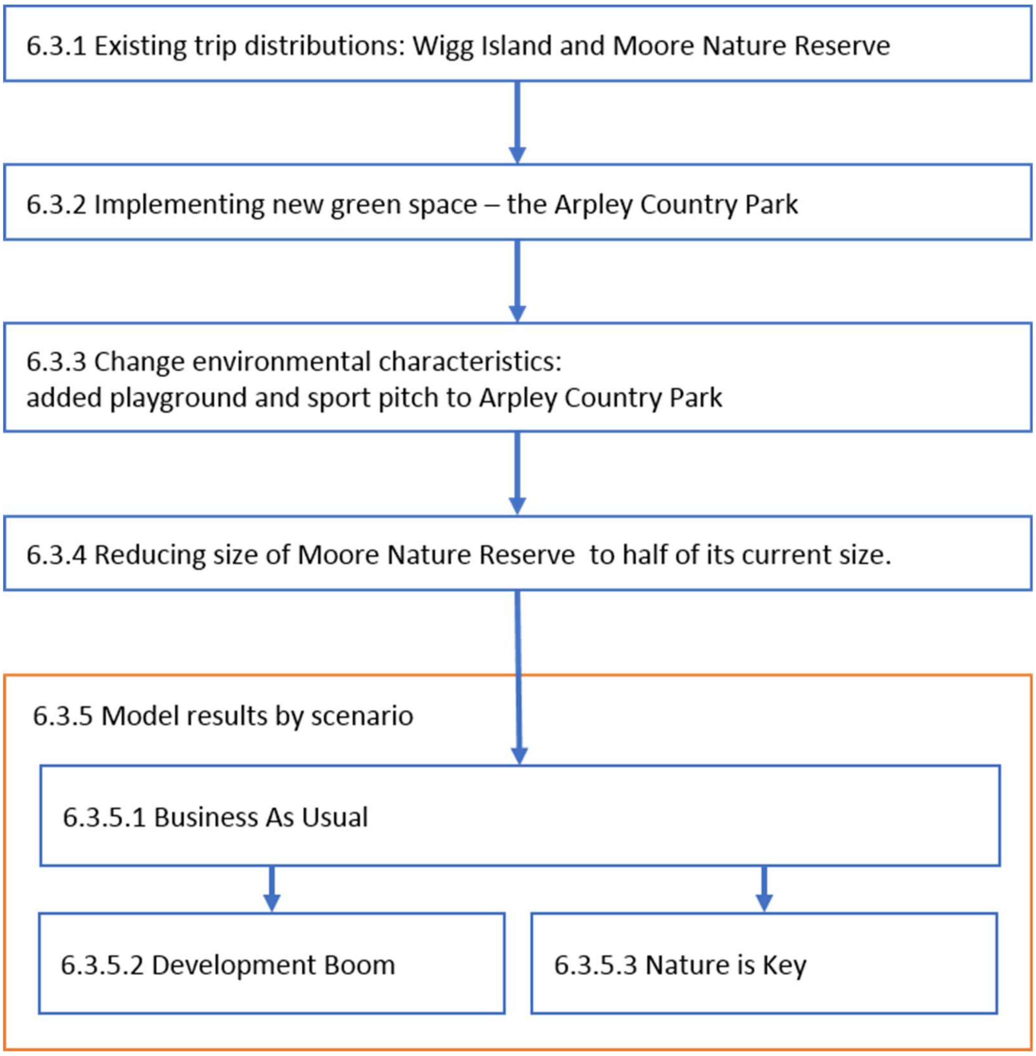


Figure 6.3. Steps of model applications.

6.3.1 Existing number of trips

The model was firstly used to estimate the existing trip distributions to Wigg Island and Moore Nature Reserve. Results are shown in Table 6.2. The trip density is calculated by dividing the total number of trips by the population of each origin on the basis of the 2011 census data. Spatial distributions of the trip densities are depicted in Figure 6.4 and Figure 6.5.

Table 6.2 *Estimated Existing Number of Trips*

	Number of trips				
	Cycling	Driving	Transit	Walking	Total
Wigg Island	5,705	20,968	721	31,810	59,474
More Nature Reserve	4,253	52,274	2,131	7,927	66,585

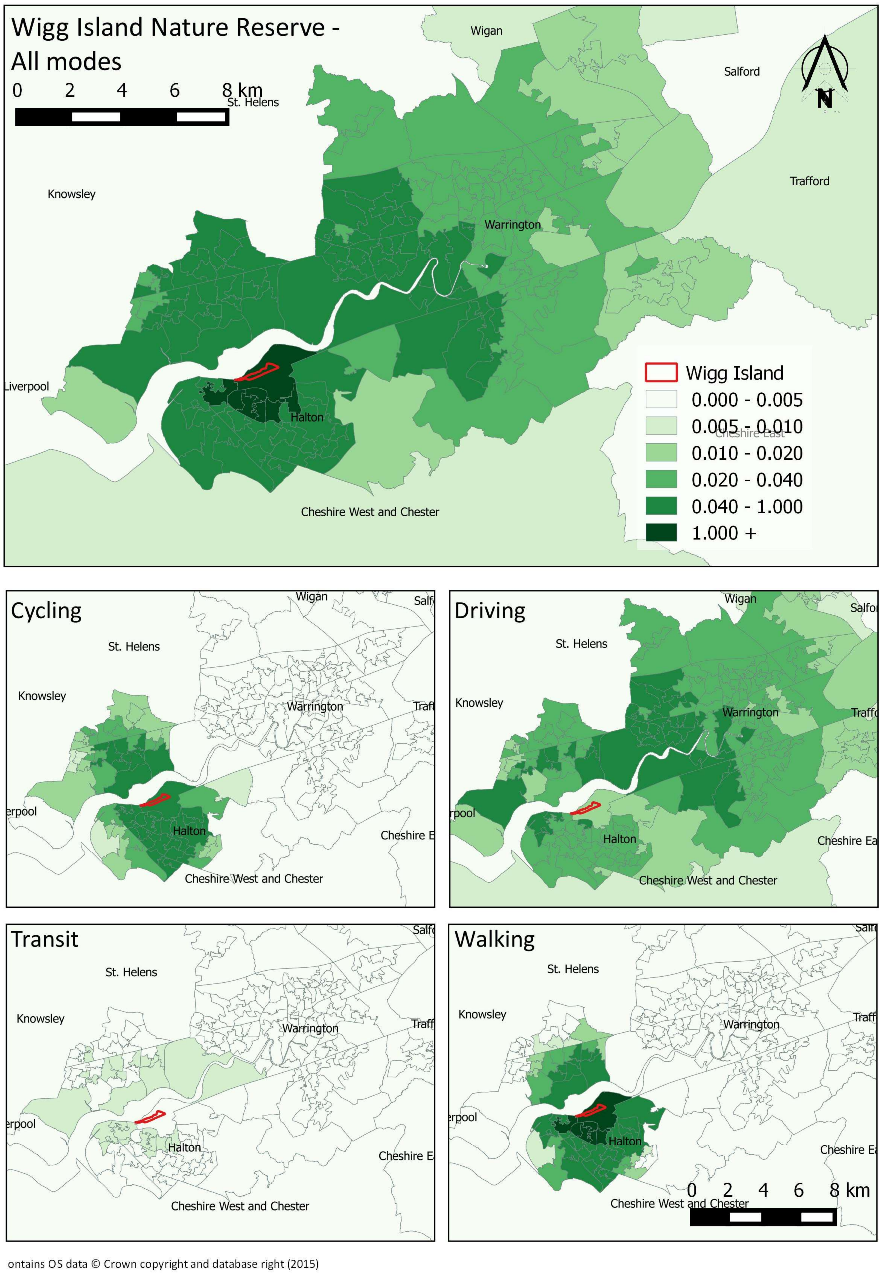


Figure 6.4. Mean of trips to Wigg Island per resident per year.

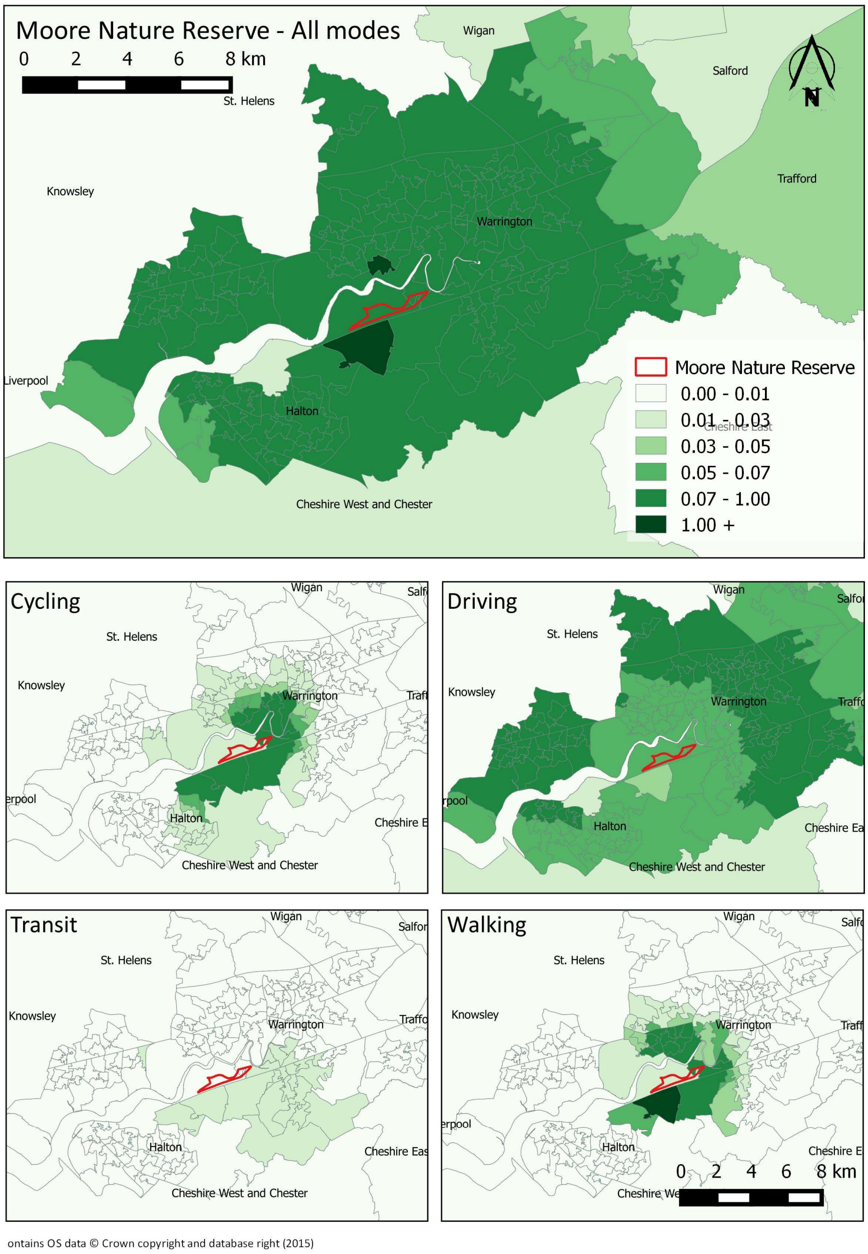


Figure 6.5. Mean of trips to Moore Nature Reserve per resident per year.

6.3.2 Implementing new green spaces—Arpley Country Park

In this test, Arpley landfill is turned into Arpley Country Park. Arpley Country Park in this section was assumed to contain features similar to Wigg Island and Moore Nature Reserve. It is an open green space with water features, woodland patches and some forms of paths; estimated results are shown in Table 6.3.

Table 6.3 *Estimated the Number of Trips When Implementing Arpley Country Park.*

	Number of trips (% changes compared with numbers in Table 6.2)				
	Cycling	Driving	Transit	Walking	Total
Wigg Island	5,658(-0.8)	20,919(-0.2)	721(-0.1)	31,809(0.0)	59,106(-0.6)
Moore	3,985(-6.3)	52,192(-0.2)	2,127(-0.2)	7,667(-3.3)	65,971(-0.9)
Arpley	6,970	70,473	2,961	20,361	100,766

As shown in Table 6.3, while Arpley Country Park is predicted to attract 100,766 trips per year; a few hundreds of trips are predicted to decrease from Wigg Island and Moore Nature Reserve if comparing the numbers under the total in Table 6.3 with Table 6.2. Since Arpley Country Park (Figure 6.6) is closer to Moore Nature Reserve (Figure 6.5), it is hardly a surprise that the effect of diverting visits from Moore Nature Reserve is more significant than for Wigg Island. Also, due to the longer distance between Arpley Country Park and Wigg Island (Figure 6.4), walking trips to Wigg Island are the least affected as shown in Table 6.3. On the other hand, walking trips to Moore Nature Reserve will be diverted by 3.3%. Cycling trips are predicted to be the most affected for both sites (0.8% decrease for Wigg Island and 6.3% decrease for Moore Nature Reserve). For longer trips, those people would travel by driving or taking public transportation, thus they are less likely to be affected (less than 0.2% decreases for both sites, and for either driving trips or transit trips).

If we compare the results in Table 6.3 with Table 6.2, regardless of transport mode, the total number of trips decreasing from Wigg Island and Moore Nature Reserve (982 in total has been decreased) is much less than the trips generated by implementing the Arpely Country Park (100,766 increase). Assuming this will be the same for any other outdoor recreational destinations in the UME area, it means transforming Arpley to a country park will encourage recreational demand significantly. Spatial distribution of trips to Arpley Country Park is shown in Figure 6.6.

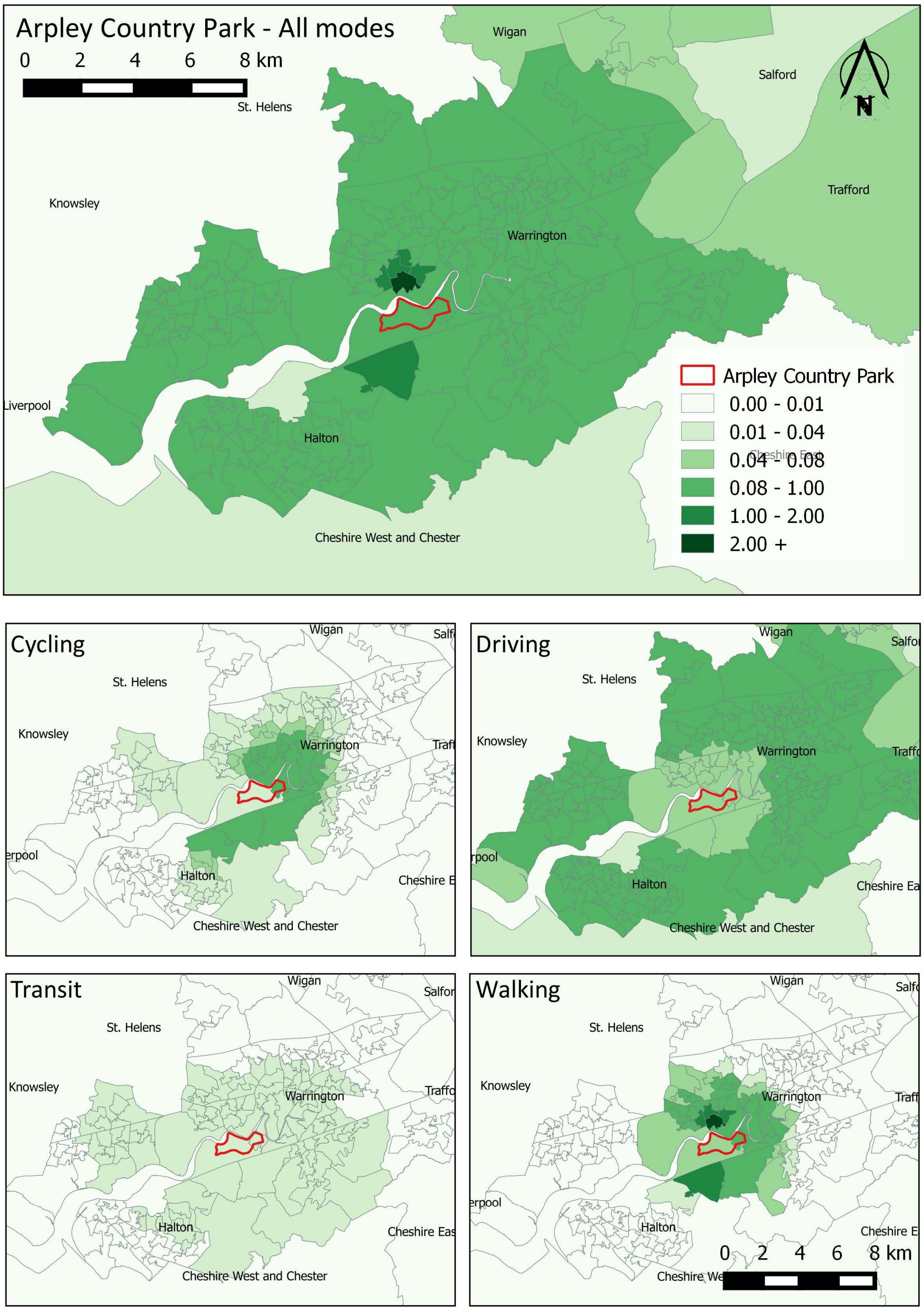


Figure 6.6.Mean of trips to the Arpely Country Park per resident per year.

6.3.3 Change characteristics of destination

According to our analysis in Chapter 4, we know people will be attracted by certain aspects of green space. The new model can estimate how much this difference will be by adding or removing feature(s). In this section, two popular features, a playground and a sports pitch, are implemented in Arpley Country Park. The changes of the travel demand to Arpely Country Park are calculated.

Table 6.4 *Estimated Number of Trips when Features are added in Arpley Country Park*

	Number of trips (% changes compared with numbers in Table 6.3)				
	Cycling	Driving	Transit	Walking	Total
Wigg Island	5,650(0.0)	20,875(-0.2)	717(-0.6)	31,808(-0.0)	59,058(-0.1)
Moore Nature Reserve	3,981(0.1)	52,095(-0.2)	2,115(-0.6)	7,462(-2.7)	65,653(-0.5)
Arpley Country Park	7,096(1.8)	153,761(118)	10,425(252)	37,074(82.1)	208,356(106)

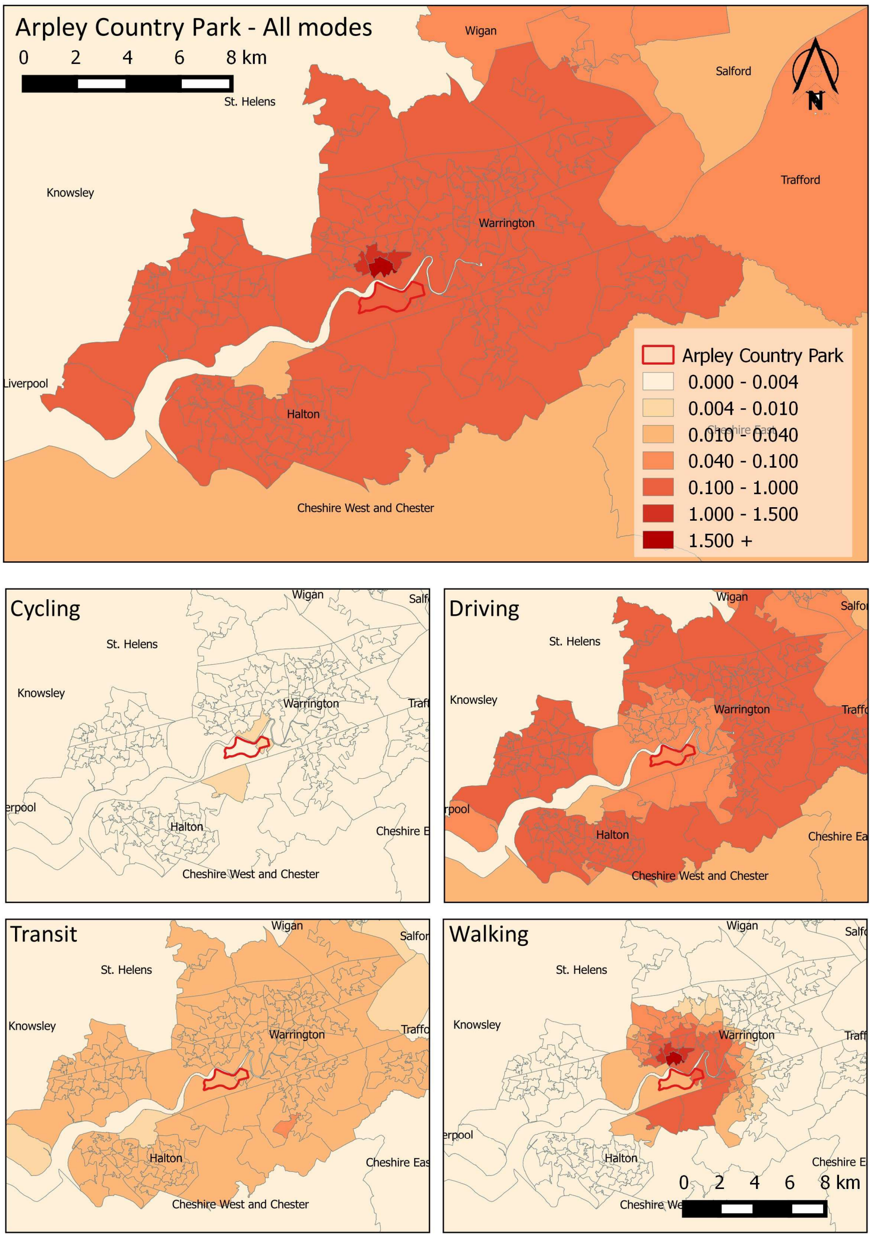
As the results reveal in Table 6.4, the total number of trips to Arpely Country Park doubled when compared with Table 6.3. The most significant change was given by transit trips (252% increase). It is two-and-a-half times larger than the estimation in Table 6.3. Increases in driving and walking trips were 118% and 82% respectively. Cycling trips only increased by 1.8%. This means when people go to the playground with their children, cycling is less favoured than any other transport modes.

If we compare Table 6.4 with Table 6.3, fewer trips will be diverted from the other destinations (366 decrease in total) compared what happened when Arpley Country Park was implemented (107,590 increase). It also means more significant outdoor recreational travel demand will be generated by adding certain facilities to the park. This is important for planning because, when space is limited, spending on improving existing green space might be more effective than implementing a new green space for attracting more outdoor recreation activities purposes.

Spatial increases of trip density (number of trips per person per year) are shown in Figure 6.7, which is calculated as:

$$\frac{D_{ja}-D_{jb}}{Pop_j} \qquad \text{Equation 38}$$

Where D_{ja} is the modelled total number of trips from Origin *j* to Arpley Country Park after adding the playground and sport field, $D_{jb}$is the modelled total number of trips before adding the playground and sport field, and Pop_j is the resident population in original *j*.



ontains OS data © Crown copyright and database right (2015)

Figure 6.7. Predicted changes in the mean of trips to Arpley Country Park per person per year.

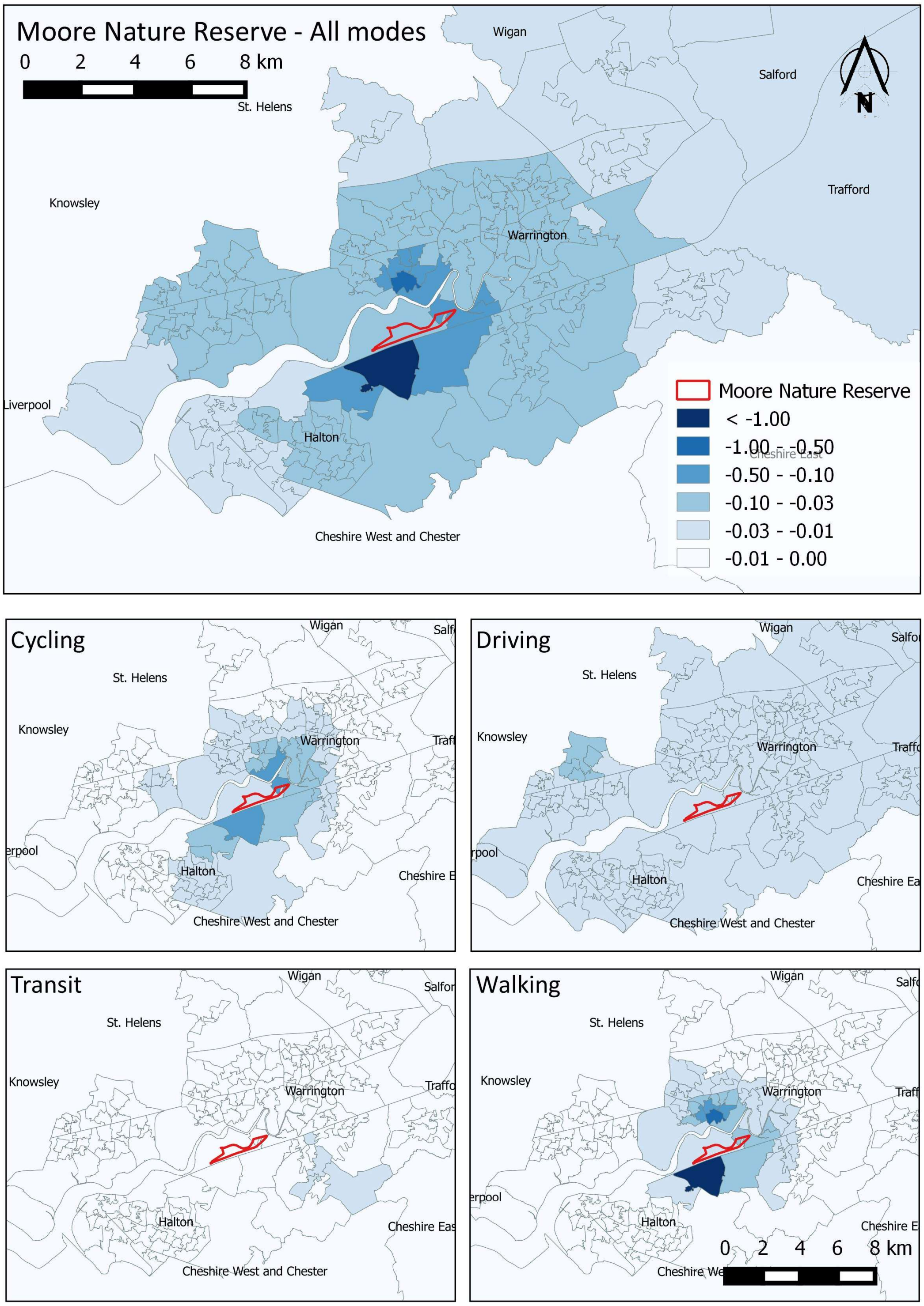
6.3.4 Reducing size of destination

In this section, the effect of size change is tested. The area of Moore Nature Reserve is cut to half of its original size as proposed by the Development Boom scenario. Compared with Table 6.4, the number of visits to the site will drop by 37.3% in total, cycling returns the most significant drop by 52.8%, and driving gives the least but still 35.2% decrease. Transit and walking trips have dropped by 41.8% and 42.3% respectively (Table 6.5). Spatial changes are shown in Figure 6.8.

Table 6.5 *Estimated Number of trips when the area of Moore Nature Reserve is reduced to by Half*

	Number of trips (% changes compared with numbers in Table 6.4)				
	Cycling	Driving	Transit	Walking	Total
Wigg Island	5,669(0.3)	20,885(0.0)	717(0.0)	31,809(0.0)	59,080(0.0)
Moore Nature Reserve	1,879(-52.8)	33,740(-35.2)	1,232(-41.8)	4,303(-42.3)	41,154(-37.3)
Arpley Country Park	7,224(1.8)	153,823(0.0)	10,431(0.1)	37,261(0.5)	208,738(0.2)

As shown in Table 6.5, there will be a significant decrease (37.3%) in travel demand for Moore Nature Reserve, and this is much bigger than the increases in Wigg Island and Arpley Country Park. These reductions are particularly significant for the short trips (walking and cycling) in the adjacent neighbourhoods as shown in Figure 6.8. Also, if we compare Figure 6.8 with Figure 6.5, the origin with the highest travel demand will be affected the most. Also, if we compare Table 6.5 with Table 6.4 for the changes in visits to Arpley Country Park, which is a very close alternative to Moore Nature Reserve, the increase (382 increase) is much less than the decrease (24,499) due to the reduction of size. Assuming same will happen to any other destinations nearby, what is very likely to happen is the outdoor recreational demand will be decreased, in particular for the nearby neighbourhoods. In other words, for some of the people who decided not to go to Moore Nature Reserve due to changes, instead of finding a replacement, they will not take an outdoor recreational trip at all.



ontains OS data © Crown copyright and database right (2015)

Figure 6.8. Predicted changes in the mean of trips to Moore Nature Reserve per resident per year.

6.3.5 Model results by scenario

In this section, the population changes are estimated and incorporated into the new model to get the finial results for each scenario. In the following sections, the results of Business as Usual BAU are calculated first as the number of trips to all three destinations. The results of BAU are then used as a baseline to demonstrate how much difference the Development Boom (DB) and Nature is Key (NK) will make.

6.3.5.1 Business as Usual (BAU)

As stated in section 6.2.2.1, this strategy aims to maintain the status quo. Developments happening in this scenario include Omega, Sandymoor and Tan House Lane. The latter will be a partial development, developed as a mix-used area as mentioned in the core strategy plan. Predicted demographic changes and related neighbourhood are shown in Table 6.6 (for details of how these numbers are calculated, please refer to Section 6.2.3).

Table 6.6 *Predicted Demographic Changes and Related Neighbourhood Based on the Business as Usual scenario*

Origin Zone (Predicted increases of population)		
Halton 003C (141)	Halton 014A (125)	Warrington 009D (1125)
Halton 007D (266)	Halton 014B (500)	Warrington 018B (625)
Halton 009A (6125)	Warrington 009B (438)	Warrington 018G (2125)
Halton 009B (2500)	Warrington 009C (1063)	Warrington 019D (116)

Modelled results are shown in Table 6.7. The characteristics of destinations are the same as those used to estimate the numbers in Table 6.4; the differences are only caused by population projections. As a result, the distributions of trips densities have not been changed from what is shown in Figure 6.4–Figure 6.6.

Table 6.7 *Predicted Number of Trips by Mode for the Business as Usual scenario*

	Cycling	Driving	Transit	Walking	Total
Wigg Island	5,807	21,255	751	32,112	59,925
Moore Nature Reserve	4,647	53,056	2,243	7,835	67,781
Arpley Country Park	8,074	156,566	11,035	38,865	214,540

6.3.5.2 Development Boom (DB)

Under this scenario, economic development is central to this strategy. There will be a full development in Omega, Sandymoor, and Tan House Lane, Port Warrington will be developed into a mix-used area, and half of Moore Nature Reserve will be taken up for the development. Predicted demographic changes and the related neighbourhood are shown in Table 6.8.

Table 6.8 *Predicted Demographic Changes and Related Neighbourhood Based on the Development Boom scenario*

Origin Zone (Predicted increases of population)		
Halton 003C (282)	Halton 014A (250)	Warrington 009D (2250)
Halton 007D (532)	Halton 014B (1000)	Warrington 018B (1250)
Halton 009A (12250)	Warrington 009B (876)	Warrington 018G (4250)
Halton 009B (5000)	Warrington 009C (2126)	Warrington 019D (232)

Table 6.9 *Predicted Number of Trips by Mode Based on the Development Boom scenario*

	Number of trips (% changes comparing with numbers in Table 6.7)				
	Cycling	Driving	Transit	Walking	Total
Wigg island	5,975(2.9)	21,645(1.8)	785(4.5)	32,416(0.9)	60,821(1.5)
Moore Nature Reserve	2,511(-46.0)	34,986(-34.1)	1,381(-38.4)	4,729(-39.6)	43,607(-35.7)
Arpley Country Park	9,231(14.3)	159,438(1.8)	11,652(5.6)	40,846(5.1)	221,167(3.1)

As results suggest in Table 6.9, the number of visits to Wigg Island and Arpley Country Park will be increased by 1.5% and 3.1% respectively. Due to the reduced size, trips to Moore Nature Reserve decreased by 35.7%, but this is 1.6% less than the number in Table 6.5 due to the increase in residents from the new developments. Spatial changes in the mean of trips to Moore Nature Reserve per resident per year are the same as that shown in Figure 6.8. The changes of total visits are shown in Figure 6.9–Figure 6.11; the positive number means an increasing number of trips by percentage compared with BAU results, and a negative number means decreasing.

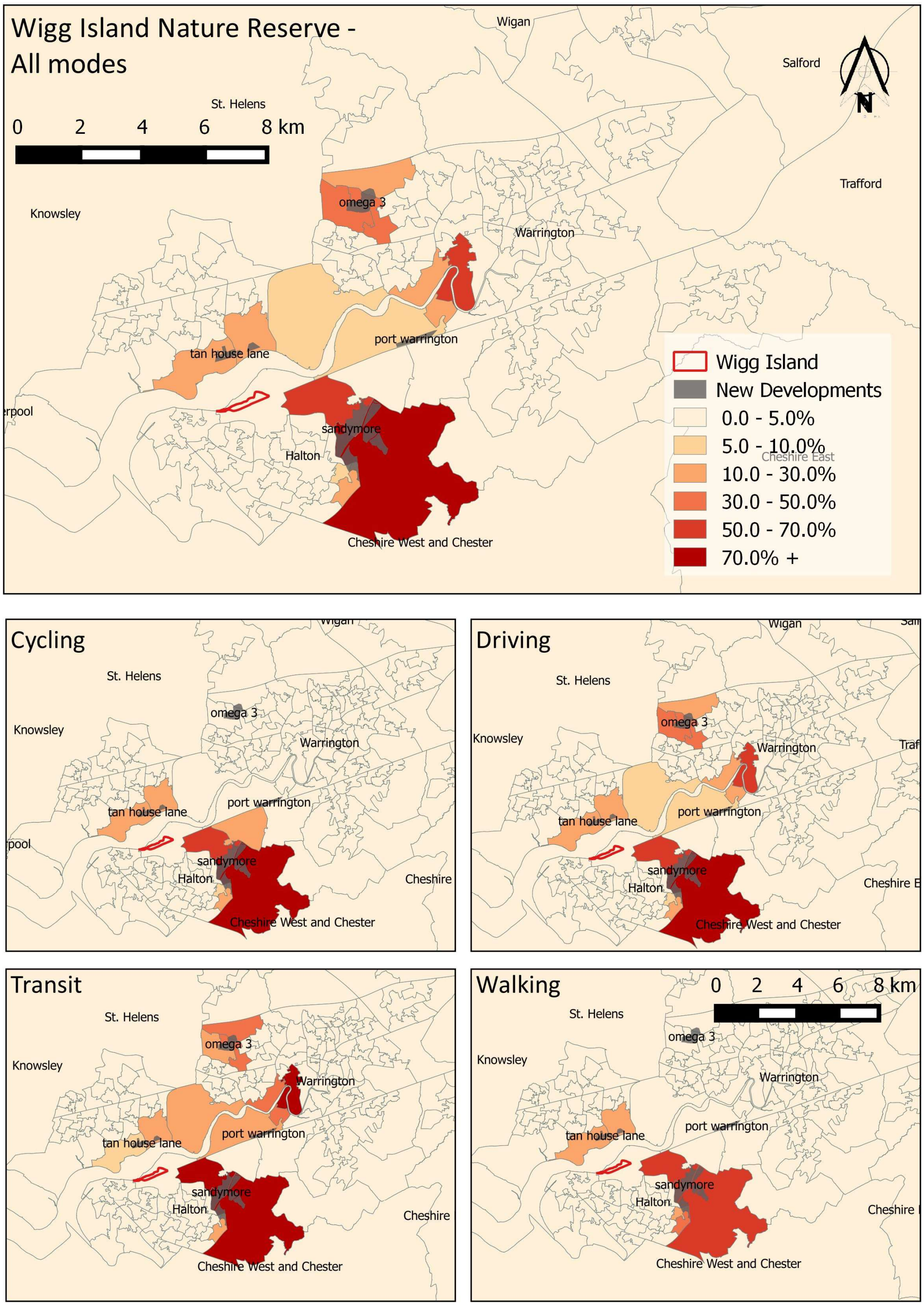


Figure 6.9. Change of the total number of trips to Wigg Island (in %) comparing Development Boom with Business as Usual scenarios.

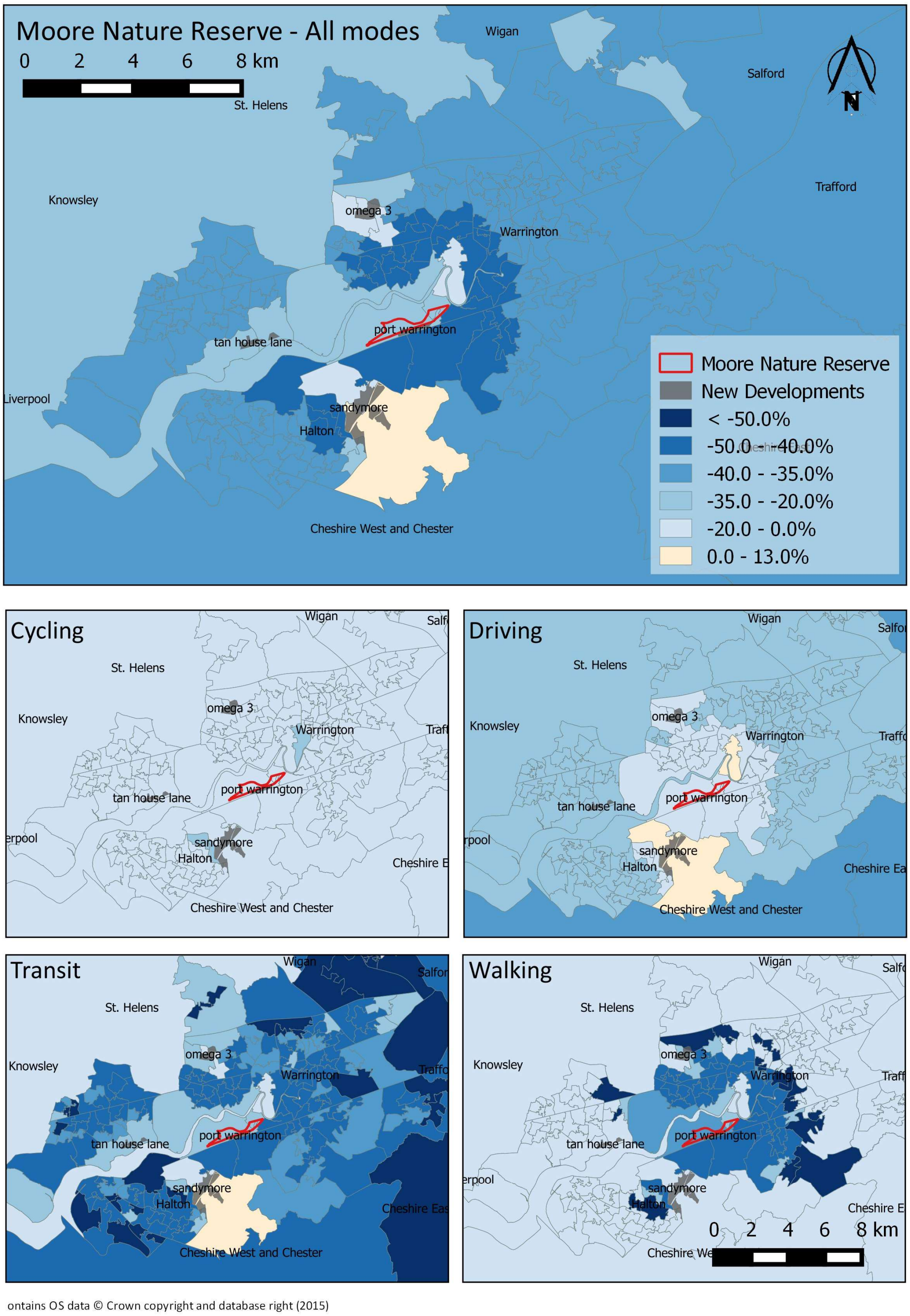


Figure 6.10. Change of the total number of trips to Moore Nature Reserve (in %) comparing Development Boom with Business as Usual scenarios.

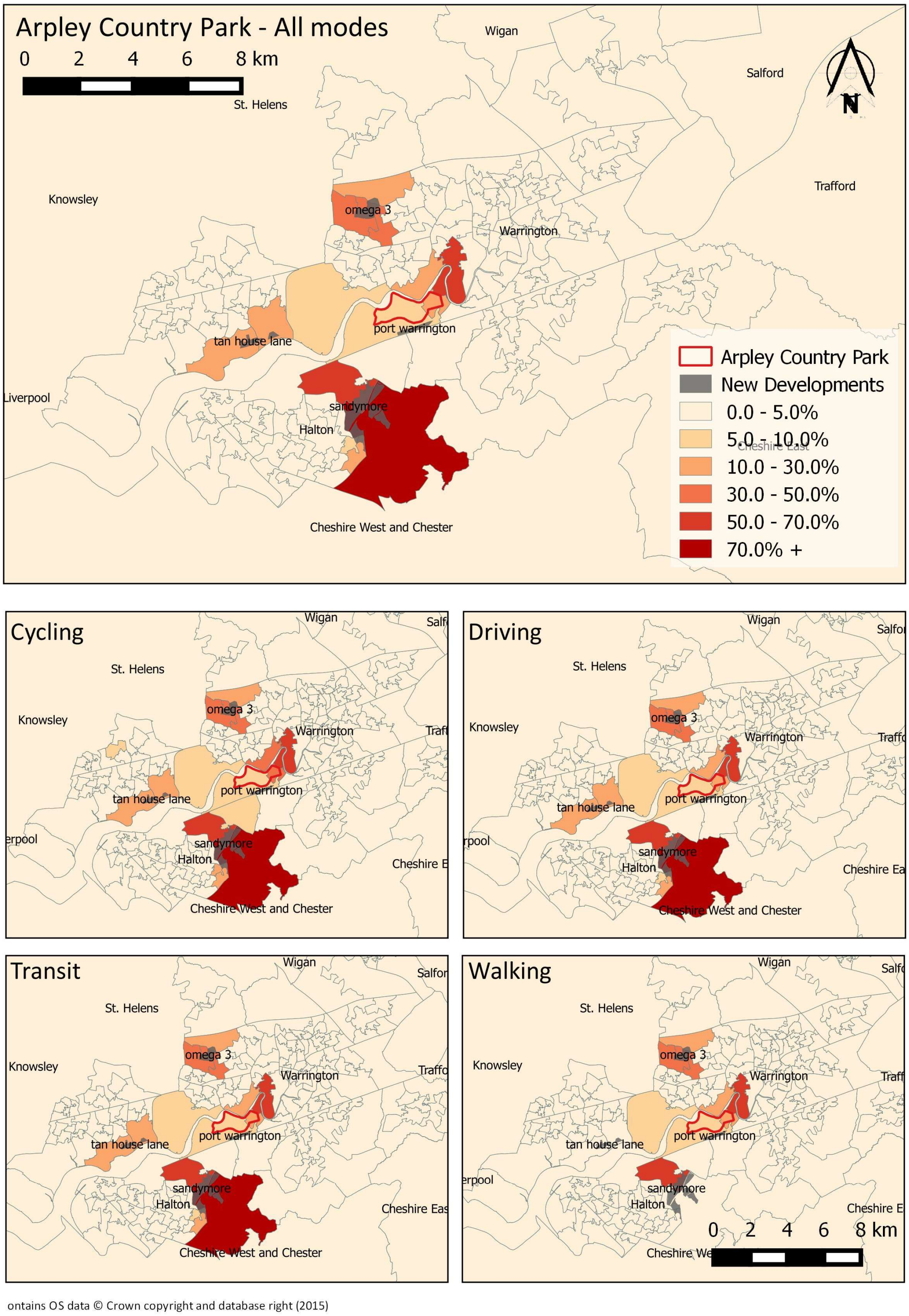


Figure 6.11. Change of the total number of trips to Arpley Country Park (in %) comparing Development Boom with Business as Usual scenarios.

As shown in maps, the increases are mainly coming from the developments proposed by the DB scenario. The developments in Tan House Lane and Sandymoore will contribute more to the increase in travel demand for Wigg Island (Figure 6.9); on the other hand, developments in Omega 3 and Port Warrington will trigger more trips to Arpley Country Park (Figure 6.11). Although the visits to Moore Nature Reserve are decreasing in general, due to the developments, outdoor recreational travel demand from the neighbourhood surrounding Sandymoore will still be increased (Figure 6.10).

6.3.5.3 Nature is Key (NK)

As the name suggests, environmental protection is the key issue in this scenario. There will be no development in Tan House Lane, partial development in Omega 3 and Sandymoore, and small-scale developments in Port Warrington without taking over any parts of Moore Nature Reserve. St Helens Canal will be improved as a wildlife site. No use of the canal for commercial or private shipping will be approved. Demographic changes under this scenario are shown in Table 6.10.

Table 6.10 *Predicted Demographic Changes and Related Neighbourhood Based on the Nature is Key scenario*

Origin Zone (Predicted increases of population)		
Halton 009A (6125)	Halton 014A (125)	Halton 009B (2500)
Halton 014B (500)	Warrington 009B (438)	Warrington 009C (1063)
Warrington 009D (563)		

Compared with the estimations under BAU (Table 6.7). There are fewer activities on all three destinations in general. This is because, firstly, the developments scales are controlled more strictly than BAU, and, secondly, the St Helens Canal is assumed to be used for wildlife and recreational purposes only. This will divert outdoor recreational travel demand towards the other three sizes.

Table 6.11 *Predicted Number of Trips by Mode Based on the Nature is Key scenario*

	Number of trips (% changes comparing with numbers in Table 6.7)				
	Cycling	Driving	Transit	Walking	Total
Wigg island	5,745(-1.1)	21,100(-0.7)	740(-1.4)	31,663(-1.4)	59,247(-1.1)
Moore	4,399(--5.3)	52,824(-0.4)	2,209(-1.5)	7,490(-4.4)	66,922(-1.3)
Arpley	7,527(-6.8)	155,925(-0.4)	10,868(-1.5)	36,324(-6.5)	210,644(-1.8)

Regarding spatial distributions as mapped in Figure 6.12–Figure 6.14, similar to the situation when Arpley Country Park was implemented (Figure 6.6), the effects are more significant on the neighbourhoods closer to the St Helen Canal and, in particular, where the recreational function is less developed.

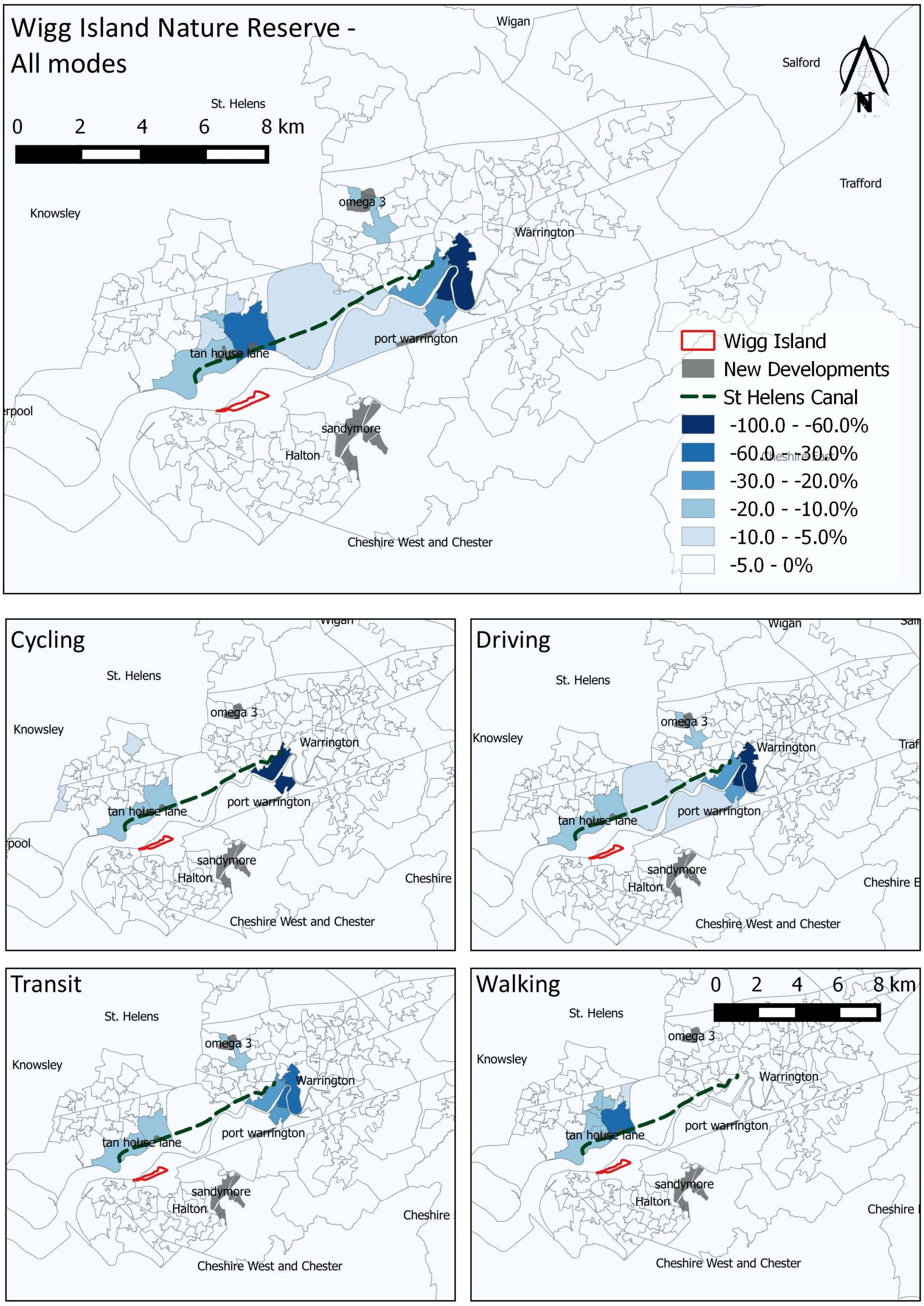


Figure 6.12. Change of the total number of trips to Wigg Island (in %) comparing Nature is key with Business as Usual.

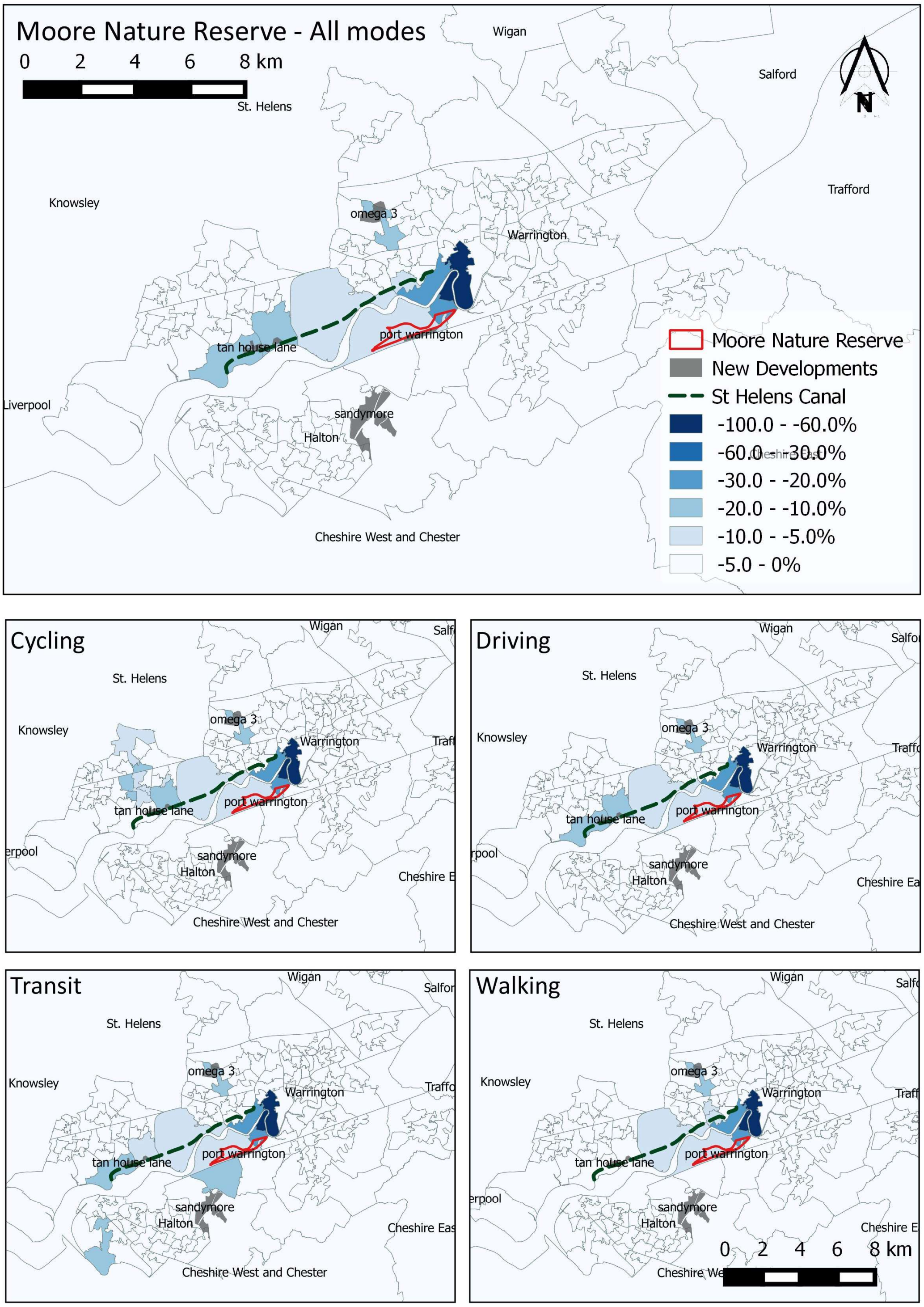


Figure 6.13. Change of the total number of trips to Moore Nature Reserve (in %) comparing Nature is key with Business as Usua.

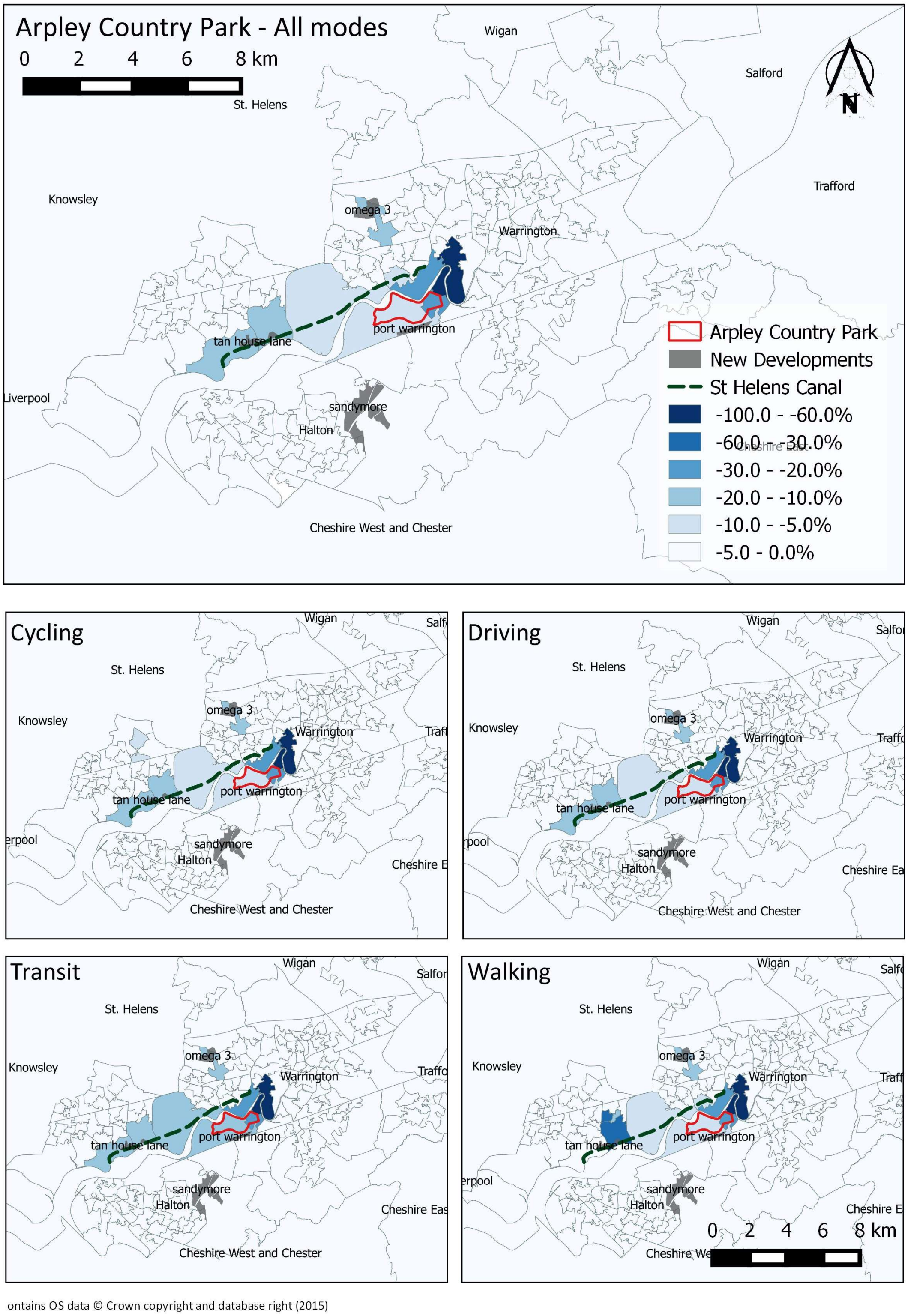


Figure 6.14. Change of the total number of trips to Arpley Country Park (in %) comparing Nature is key with Business as Usua.

6.4 Summary

In this chapter, the new model was applied to estimate changes in outdoor recreational travel demand to Wigg Island, Moore Nature Reserve, and Arpley Country Park. Those changes were trigged by land use and population differences. The following findings can be summarised from the tests in this chapter.

Firstly, introducing a new destination (i.e., Arpley Country Park) diverts trips from existing green spaces, and this impact was more significant on Moore Nature Reserve, which is closer to Arpley Country Park, and walking and cycling trips were more affected than driving and transit trips (Section 6.3.2).

Secondly, installing a playground and sports pitch in Arpley Country Park significantly increased visits to the green space (Section 6.3.3). This may mean a lack of playground and sports pitches in neighbourhoods needs further investigations, but the travel demand was clearly increased when new green space is implemented, and the demand will be boosted even further when new features are added to the site.

Thirdly, the size of green space significantly affects the number of visits (Section 6.3.4), and, as expected, the change of travel demand was in the same direction as the changes of scale. Among the four travel modes, the cycling trip was affected the most; driving trip was the least.

Fourthly, the proximity of residential areas to green space is a double-edged sword. On the one hand, the proximity of residential areas can make green space more popular. On the other hand, it could also increase the risk of over use, leading to disturbances of wildlife habitats and degradation of the green space. This was clearly demonstrated by the results of testing three scenarios through the new model (Section 6.3.5).

Also, as results the show in Section 6.3.5, compared with the BAU scenario (Table 6.7), DB gave significant increases in travel demand to Wigg Island and Arpley Country Park (

Table 6.9). Also, in the DB scenario, part of Moore Nature Reserve was replaced by development, and travel demand to Moore Nature Reserve was decreased in general. However, as Figure 6.10 shows, there is still an increase in the neighbourhoods with proposed development (i.e., Sandymoore). Regarding the NK scenario, it returned only decreases due to the limitations put on developments (Table 6.11).

Chapter 7 Discussion and Conclusion

7.1 Discussion

In this research, a new travel demand model for outdoor recreational trips has been developed, using a method from mainstream transport forecasting modelling—the Random Utility Maximisation (RUM) theory-based Discrete Choice Models (DCMs). This is to fill the gap in our knowledge in applying convectional transport forecasting method to analyse outdoor recreational trips as reviewed in Chapter 2. In order to fill the knowledge gap, four research questions were answered during this research which will be addressed in the following four sections of this chapter.

7.1.1 How to build the new model for outdoor recreational travel?

As reviewed in Chapter 2, conventional transport modelling has been developed since the 1950s (Boyce & Williams, 2015). However, it has rarely touched the field of outdoor recreational trips. This is because transport appraisal is usually focused on estimating reductions in congestion and the benefits generated by travel-time saving (Hollander, 2016). Outdoor recreational trips, on the other hand, normally happen during the off-peak time when congestion is less likely to occur. A group of environmental economics researchers have been doing studies on modelling outdoor recreational activities. However, the conventional transport forecasting method – the Random Utility Maximisation (RUM) based Discrete Choice Methods (DCMs) have been mainly applied to a single activity/habitat (Phaneuf & Smith, 2005). This has made it difficult to apply the results to either different places from the case study areas or for general planning practices. Attention, however, to the outdoor recreational trips has increased over the last decade.

Since 2009, Natural England has funded DEFRA (Department for Environment, Food & Rural Affairs) and the Forestry Commission to conduct a survey named Monitor of Engagement with the Nature Environment (MENE). This survey has provided robust evidence for studying the travel demand for outdoor recreational trips. To date, the only travel demand model built on this set of data has been completed by Sen et al. (2012, 2014). However, it is developed through different modelling theory—the Negative Binomial Regression (NBR) for applications at the national and regional scale. This model has been reviewed in Section 2.6, and it is acknowledged that it has a weak power in estimating the number of outdoor recreational trips to single destination.

Therefore, the new travel demand model follows a conventional transport modelling structure. As shown in Figure 4.21, the new model started from the trip-generation function. The trip-generation function estimates a total number of trips generated from each origin, calculated by multiplying the mean of outdoor recreational trips per person per year by the population of each origin. The second step is the modal split function in this particular model. It is in the form of a multinomial logit model, which splits the total number of trips generated from each origin into one of the four transport modes: cycling, driving, transit (public transport) and walking. Finally, another group of the multinomial logit models has been used as the trips distribution functions, dividing the number of trips further to all outdoor recreational destinations. The outcome of the new model is the total number of visits to any outdoor recreation destination from all origins.

This new model is built upon reviews of previous studies, as well as various experiments which were demostrated in this research. The variables used in this research are drawn from previous studies as reviewed in Section 4.3. These include travel time, land uses, land cover, population, the percentage of retired people and non-white ethnicity group, income, travel distance, and activities. Unlike the previous studies, the tests in this research have found an insignificant correlation between outdoor recreational trips generation and social demographic variables, i.e., the percentage of retired people, non-white ethnicity group and income. Therefore, the population is the only variable kept in the trip-generation function of this research.

The next calibration process began with finding out the best model structure. Three different forms of DCMs were tested. Two main findings from the results are, firstly, for outdoor recreational travel demand forecasting, the transport mode choice is decided before the selection of destination. This result is aligned with Rohr’s (2005) previous findings. Secondly, regarding the trip-distribution functions, travel time needs to be considered as well as the characteristics of the sites. In the existing transport trip distribution models, travel time and generalised cost are normally the only variables included (Ortúzar & Willumsen, 2011). Environmental characteristics of destinations are usually missing in transport modelling.

Before the model form was finalised, three more tests have been carried out to investigate the best combination of variables. Based on the data available to this research, the experiments suggested that grouping individuals by either travel distance or travel activity play an insignificant role in improving the model. Since there are no previous studies regarding this topic, it is not certain that this result is not caused by the samples size this reseearch has had for each group. However, the mixed logit model tests carried out in section 4.6 suggest that there are clear variations among people who decided to visit either mountains or coastlines. It means it is worth carrying out further studies separately. Also, the final logit models assume a precise sequence of decisions (i.e. first travel mode, then destination). There are also variations among individuals.

7.1.2 Is the estimation accurate enough?

This question is answered through the first part of the validation process. The first group of validations was done in two outdoor recreational destinations (Wigg Island and Wigan Flashes) in the model calibration area. The results calculated by the new model were compared with on-site observations from Wigg Island and Wigan Flashes. As the results shown in Tables 5.1–5.3, the new model provides a robust estimation on Wigg Island with only a 0.5% difference when compared with the data collected from the people-counting monitor. Regarding Wigan Flashes, the total number of visits was 24% higher than the estimation made by the manager from Wigan Flashes. There is no evidence to prove that the estimation made by the manager is acurate. And we know model results have had a certain level of uncertainty as well. Therefore, this relatively difference is considered to be caused by both. However, if the difference between the model result and observation was big, as a standard procedure of travel-demand modelling (Hollander, 2016), an attraction residual (-0.22) was added (the number was calibrated by experiments), and the difference was reduced to 0.7%. In conclusion, the new model can make robust estimations in the model calibrating area.

7.1.3 To what extent can the new model be transferred to destinations outside the case study area?

The next group of validations was done on the ten English National Parks. In order to explore the answer to the question regarding the extent to which this new model can be transferred to destinations outside the model calibaration study area, the estimated number of visits to each English National Park was compared with the amount calculated through the Scarborough Tourism Economic Activity Monitor (STEAM) model and reported in the National Parks' report. As the results showed in Tables 5.5–5.6, the new model has made very accurate (less than 10% difference) estimations for the parks which are close to the model calibration area, except the Lake District, where the model hugely underestimated the visits. The model was built upon the sampled observations of outdoor recreational travels in the North-West region. The popularity of the Lake District is extraordinary, it is not surprising that, the model did not capture the speciality of the Lake District. For the national parks which are far away from the model area, the differences are between 10% to 20%, with only one exception, Exmoor National Park. One cause of this difference is that, the new model built in this research is based on the behaviour of individuals from the North-West region. Residents in the south might behave differently, even towards the same land use feature. One solution to solve above problem is to build separate models for different area. The method used in this research can be readily applied to the data specifically selected for the targeting area.

Although this new model has very strong transferability, it has shown unstable performance in estimating the site in the south of England when applied to different regions. The practical findings regarding peoples' behaviour towards outdoor recreational trips are only relevant to the North-West region of UK. In order to minimise bias caused by major behaviour variations, it is necessary to calibrate the model via different datasets selected by targeting geographic areas whenever the model is applied to an area further away from the North-West region. As mentioned above, all the data used in this study are available online and the same data is publicly available for other sites. Therefore, the same method can be applied to any place in the UK. The method can also be used in any other countries where similar data are available.

One primary challenge in this research is to obtain detailed observations for validation. It is rare to find origin-destination surveys or even systematic arrivals for greenspace sites. Therefore, it might be challenging to find validation data for different sites in the UK, let alone for other countries. New social media data may offer some potential in filling this gap.

It would, thus, seem necessary to develop a new research agenda to collect such information to strengthen the empirical basis of the model.

7.1.4 How can city planners and designers use this new method?

Given that the new model has been proven to be capable of making robust estimations in the calibration area, in the last part of this research, Chapter 6, it is demonstrated how decision makers can make use of this new model to test scenarios for change. The scenario tests were done on the Upper Mersey Estuary area, inside the model calibration area. Three scenarios were tested (i.e., Business as Usual, Development Boom, and Nature is Key), and the results indicate that spatial locations and the environmental characteristics of green spaces are essential for outdoor recreation demand, weighted differently by people who would choose different transport mode. Urban planners can use this new model to assess how much the visits to the green space will change by implementing different strategies, thereby anticipating any outcomes related to these changes.

From a planning point of view, the research reported here has been informed by previous studies that demonstrate that many characteristics of the outdoor recreational sites have played essential roles in attracting visitors (e.g., Jones et al., 2010; Paracchini et al., 2014; Sen et al., 2014). What is not in the existing literature is that individuals who choose different transport mode would weigh the characteristics differently. From this research, we can learn that cyclists tend to go to the country parks more often than to the parks in cities. The contrary has been found for people who travelled by public transport. The playground appears to be an important feature to everyone (in particular for individuals who choose to drive and walk) except people who would like to cycle. Therefore, when planning/designing a green infrastructure, it is essential to vary its features for different users. This model can be used to estimate who will be the most likely group of people to use the selected destination, enabling stakeholders to plan accordingly.

Secondly, compared with adding new green space, improving the existing outdoor recreation destinations by implementing new features can have a similar if not a greater effect on increasing residents’ outdoor recreation demands. This is especially important when space and funding are limited. Moreover, the impact of reducing the size of green spaces must be carefully studied because it might depress demand for outdoor recreation dramatically. This consequence might be challenging to mitigate by installing a different green space somewhere else, particularly for those who have been affected. Furthermore, there are potential crises by building a new green space as well. For example, the new green space will compete with exiting green spaces. Hence it could reduce the trips to currently less-visited green spaces, potentially raising the likelihood of crime and vandalism.

7.2 Conclusion

The unique feature of this research is that the new model provides the first quantitative insights into the effects on green spaces resulting from planning and design decisions of location, size, land use, environmental characteristics and transport connections. There is no similar kind of model existing in the transportation field as reviewed in Chapter 2. The most relevant model is the Negative Binomial Regression (NBR) model from the environmental economics area, which was developed by Sen et al. (2013,2014). The application of NBR to the UK NEA project indicates that such a tool is very useful for urban planning and landscape planning. However, the NBR model has had significant weakness compared with the DCMs, given its inconsistent performance both in theory and in practice (Ortúzar & Willumsen, 2011). Apart from that, the NBR outdoor recreation model developed by the UK NEA project is only for assessments at the national/regional scale, not for the individual site (as noted in Section 2.6). In contrast with the existing NBR model, the new DCMs-based models can predict the number of trips to any green space. This new model developed through this research provides a new method for assessing the impacts of alternative urban development and green space designs.

Secondly, the new model has a very systematic and rigorous calibration and validation process. The new model is built upon reviews of previous studies as well as experiments, wherein three different forms of DCMs have been designed and tested on an expanded database. During the model-calibaration process, this research has incorporated, for the first time, a wide range of data in modelling trips to green spaces, establishing entirely new methods for forecasting travel to green spaces by combining data sources on transport, census, land use and natural environment. All of these data are published data, enabling this method to be transferred to any site in England easily. Regarding the validation process, the new model was tested on independently observed data that had not been used in model calibration, unlike what would normally be done during transport modelling: divide the same dataset into two parts, one for calibration and another for validation (Hollander, 2016). As a result, the new model has been validated more rigorously, and it has given robust estimations, in particular in those places close to the model-calibration area.

From the application point of view, the new model has been applied to alternative land use and green space scenarios, and it provides new information to the valuation of green spaces. It provides empirical evidence of how much the outdoor recreational demand will be affected by adding new green space and changing characteristics of green space. Depending on how the new model is applied, it could provide guidance on how green space should be designed and located to obtain greater health and well-being gains for the population.

References

Bateman, I. J., Abson, D., Beaumont, N., Darnell, A., Fezzi, C., Hanleys, N., … Termansen, M. (2011). Economic values from ecosystems. *Human Well-Being, UK National Ecosystem Assessment: Technical Report, 81*, 1067–1152.

Bateman, I. J., Harwood, A. R., Abson, D. J., Andrews, B., Crowe, A., Dugdale, S., ... & Hulme, M. (2014). Economic analysis for the UK national ecosystem assessment: synthesis and scenario valuation of changes in ecosystem services. Environmental and Resource Economics, 57(2), 273-297.

Bates, J. J., Williams, I., Coomber, D. & Leather, J. (1996) The London congestion charging research programme: 4. The transport models. *Traffic Engineering and Control 37*, 334-339.

Ben-Akiva, M. E. (1973). *Structure of passenger travel demand models* (doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Bly, P., Emmerson, P., Paulley, N. & Van Vuren, T. (2002). *User-friendly multi-stage modelling advice. Phase 3: multi-stage modelling options*. Crowthorne, Berkshire: Transport Research Laboratory.

Bowker, J. M., Bergstrom, J. C. & Gill, J. (2007). Estimating the economic value and impacts of recreational trails: A case study of the Virginia Creeper Rail Trail. *Tourism Economics, 13*(2), 241-260.

Boyce, D. E., & Williams, H. C. (2015). Forecasting Urban Travel: Past, Present and Future. Edward Elgar Publishing.

Boyd, J., & Mellman, J. (1980). The effect of fuel economy standards on the U.S. automotive market: A hedonic demand analysis. *Transportation Research 14*, 367–378.

Bruton, M. J. (1975). *Introduction to transportation planning* (2nd ed.). London: Hutchinson.

Cameron, A. C., & Trivedi, P. K. (1990). Regression-based tests for overdispersion in the Poisson model. *Journal of Econometrics, 46*, 347-364.

Cardell, S., & Dunbar, F. (1980). Measuring the societal impacts of automobile downsizing. *Transportation Research 14*, 423–434.

Daly, A. (2013) Forecasting behaviour: with applications to transport. In S. Hess and A. Daly (eds.), *Choice modelling* (pp. 48-72). Cheltenham, UK, and Northampton, MA: Edward Elgar Publishing.

Daly, A. J. (1997). Improved methods for trip generation. *Proceedings of the 25th European Transport Forum*, Brunel University, September 1997, England.

Daly, A. J., & Zachary, S. (1978). Improved multiple choice models. In D.A. Hensher and M.Q. Dalvi (eds.), *Identifying and measuring the determinants of modal choice* (pp. 335-357). London: Saxon House.

Davies, L., Kwiatkowski, L., Gaston, K. J., Beck, H., Brett, H., Batty, M., ... & Perino, G. (2011). Urban. In The UK national ecosystem assessment: technical report. UNEP-WCMC.

Domencich, T., & McFadden, D. (1975). Urban travel demand: a behavioural approach. Amsterdam: North-Hollan Publishing Co.

Douglas, A. A., & Lewis, R. J. (1970). Trip generation techniques: (1) Introduction; (2) Zonal least squares regression analysis. *Traffic Engineering and Control, 12*, 362–365, 428–431.

Feather, P. (1994). Sampling and aggregation issues in random utility model estimation. *American Journal of Agricultural Economics, 76*(4), 772–780.

Fratar, T. J. (1954). Vehicular trip distribution by successive approximations. *Traffic Quarterly,* 8(1).

Friedrich, M., Mott, P. & Nökel, K. (2000). Keeping passenger surveys up to date, *Transportation Research Record, 1735*, 35-42.

Fuller, R. A., Irvine, K. N., Devine-Wright, P., Warren, P. H. & Gaston, K. J. (2007). Psychological benefits of green space increase with biodiversity. *Biology Letters, 3*(4), 390-394.

Hall, M. D., Van Vliet, D. & Willumsen L. G. (1980). SATURN- a simulation-assignment model for the evaluation of traffic management schemes. *Traffic Engineering and Control, 21*, 168-176.

Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S. & Gärling, T., 2003. Tracking restoration in natural and urban field settings. *Journal of Environmental Psychology, 23*, 109–123.

Hicks, R., & Strand, I. (2000). The extent of information: Its relevance for random utility models. *Land Economics, 76*(3), 374–385.

Hollander, Y. (2016). *Transport modelling for a complete beginner*. CTthink!.

Hutchinson, B. G. (1974). *Principles of urban transport systems planning*. New York: McGraw-Hill.

Jahanshahi, K., Williams, I., & Hao, X. (2009). Understanding travel behaviour and factors affecting trip rates. In *European Transport Conference.* Leeuwenhorstm, the Netherlands.

Jiao, X., Jin, Y., Gunawan, O. & James, P. (2015). Modelling spatial distribution of outdoor recreation trips of urban residents. *International Review for Spatial Planning and Sustainable Development, 3*(3), 36-49.

Jones, A., Wright, J., Bateman, I. & Schaafsma, M. (2010). Estimating arrival numbers for informal recreation: A geographical approach and case study of British woodlands. *Sustainability, 2*(2), 684–701.

Jones, C., & Lupi, F. (1997). *The effect of modeling substitute activities on recreation benefit estimates.* Working Paper. Damage Assessment Center, National Oceanic and Atmospheric Administration.

Kling, C. L., & Thomson, C. J. (1996). The implications for model specification for welfare estimation in nested logit models. *American Journal of Agricultural Economics, 78,* 103–114.

Kling, C. L., & Thomson, C. J. (1996). The implications for model specification for welfare estimation in nested logit models. *American Journal of Agricultural Economics, 78*, 103–114.

Lave, C. A. (1969). A behavioural approach to modal split forecasting. *Transportation Research, 3,* 463-480.

Lisco, T. E. (1967). *The value of commuters' travel time: A study in urban transportation* (doctoral dissertation). University of Chicago.

Luce, R. D. (1959). *Individual choice behaviour*. New York: Wiley.

Lynch, J. T. (1959). Home-interview surveys and related research activities. *Public Roads, 30,* 185–186.

Martinez-Espineira, R., & Amoako-Tuffour, J. (2008). Recreation demand analysis under truncation, overdispersion, and endogenous stratification: An application to Gros Morne National Park. *Journal of Environmental Management, 88*(4), 1320-1332.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. In P. Zarembka (ed.), *Frontiers in Econometrics*. 105-142.

McFadden, D. (1976/1997). The theory and practice of disaggregate demand forecasting for various modes of urban transportation. In T.H. Oum et al. (eds.) *Transport economics* (pp. 51-80). Amsterdam: Harwood.

McFadden, D. (1978). Modelling the choice of residential location. In A. Karlqvist, L. Lundqvist, F. Snickars and J. Weibull (eds). Spatial interaction theory and residential location (pp. 75-96). Amsterdam: North-Holland.

McFadden, D. (2001). Disaggregate behavioural travel demand's RUM side: A 30-year retrospective. In D. Hensher (ed.). Travel behaviour research (pp. 17-63). Oxford: Pergamon.

McFadden, D., & Train, K. (2000). Mixed MNL models of discrete response. *Journal of Applied Econometrics, 15*, 447–470.

McGillivray, R. G. (1970). Demand and choice models of mode split. *Journal of Transport Economics and Policy, 4,* 192-207.

Mitchell, R. B., & Rapkin, C. (1954). *Urban traffic*. New York: Columbia University Press.

Miyagi, T. (1984). *The conjugate dual approach to travel demand modelling*. PTRC 125h Summer Annual Meeting.

MVA (2005). *Multi-modal model data provision. Report for the Denvil Coombe Practice on behalf of the Integrated Transport and Economic Appraisal Division*. London: UK Department for Transport.

Natural England (2016). *Monitor of engagement with the Natural Environment Spatial Report 2015 – 2016*. Available from https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/614353/mene-headline-report-2015-16.pdf

Ortúzar S., & Willumsen, L. (2011). *Modelling transport* (4th ed.). Chichester: Wiley.

Parsons, G., & Kealy, M. J. (1992). Randomly drawn opportunity sets in a random utility model of lake recreation. *Land Economics, 68*(1), 93–106.

Parsons, G., & Kealy, M. J. (1995). A demand theory for number of trips in a random utility model of recreation. *Journal of Environmental Economics and Management, 29*(3), 357–367.

Parsons, G., Massey, D. M., & Tomasi, T. (2000). Familiar and favorite sites in a random utility model of beach recreation. *Marine Resource Economics, 14*, 299–315.

Parsons, G., Plantinga, A., & Boyle, K. (2000). Narrow choice sets in a random utility model of recreation demand. *Land Economics,* 76(1), 86–99.

Parsons, G. R., & Hauber, A. B. (1998). Spatial boundaries and choice set definition in a random utility model of recreation demand. *Land Economics, 74* (February), 32–48.

Perino, G., Andrews, B., Kontoleon, A., & Bateman, I. (2011). *Urban greenspace amenity-economic assessment of ecosystem services provided by UK urban habitats*. Report to the UK National Ecosystem Assessment, University of East Anglia, Norwich. [online] Available from http://uknea. unep-wcmc. org/Resources/tabid/82/Default. aspx

Peters, T., Adamowicz, W., & Boxall, P. (1995). Influence of choice set considerations in modeling the benefits from improved water quality. *Water Resources Research, 31*(7), 1781–1787.

Phaneuf, D. J., & Smith, V. K. (2005). Recreation demand models. *Handbook of Environmental Economics, 2,* 671-761.

Quarmby, D. A. (1967) Choice of travel mode for the journey to work. *Journal of Transport Economics and Policy, 1,* 273-314.

Revelt, D., & Train, K. (1998). Mixed logit with repeated choices. *Review of Economics and Statistics, 80,* 647–657.

Roberts, M., & Simmonds, D.C. (1997). A strategic modelling approach for urban transport policy development. *Traffic Engineering and Control, 38,* 337-384.

Rohr, C. (2005). *The PRISM model: Evidence on model hierarchy and parameter values. Report for the UK Department for Transport*. Cambridge: RAND Europe.

Sen, A., Darnell, A., Crowe, A., Bateman, I. J., & Munday, P. (2012). *Economic assessment of the value of open-access recreation in UK ecosystems: A scenario analysis.* Centre for Social and Economic Research on the Global Environment (CSERGE), School of Environmental Sciences, University of East Anglia.

Sen, A., Harwood, A. R., Bateman, I. J., Munday, P., Crowe, A., Brander, L., … Provins, A. (2014). Economic assessment of the recreational value of ecosystems: Methodological development and national and local application. *Environmental and Resource Economics, 57*(2), 233–249.

Shaw, W. D., & Ozog, M. T. (1999). Modeling overnight recreation trip choice: application of a repeated nested multinomial logit model. Environmental and Resource Economics, 13(4), 397-414.

Shrestha, R. K., Stein, T. V., & Clark, J. (2007). Valuing nature-based recreation in public natural areas of the Apalachicola River region, Florida. *Journal of Environmental Management, 85*(4), 977-985.

Shuldiner, A. T., & Shuldiner, P. W. (2013). The measure of all things: reflections on changing conceptions of the individual in travel demand modelling. Transportation, 40(6), 1117-1131.

Smith, L., Beckman, R., Baggerly, K., Anson, D. & Williams, M. (1995) *TRANSIMS: Transportation Analysis and SIMulation System, Project Summary and Status*. A report by the Los Alamos National Laboratory to US Department of Transportation and US Environmental Protection Agency, Washington, DC. Available from ntl.bts.gov/DOCS/446/html.

Spear, B. D. (1977). *Applications of new travel demand forecasting techniques to transportation: A study of individual choice models*. Final Report to the Office of Highway Planning, Federal Highway Administration, US Department of Transportation, Washington, DC.

Stopher, P. R. (1969). A probability model of travel mode choice for the work journey. *Highway Research Record, 283*, 57-65.

Stopher, P. R., & Meyburg, A. H. (1975). Urban transportation modeling and planning. Lexington, MA: Lexington Books.

Supernak, J. (1983). Transportation modelling: Lessons from the past and tasks for the future. *Transportation, 12*, 79–90.

Taylor, N.V. (2003). The CONTRAM dynamic traffic assignment model. *Networks and Spatial Economics, 3,* 297-322.

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

UN Habitat (2016). Habitat, U. N. (2016). Urbanization and Development Emerging Futures. World Cities Report.

Van Vuren, T. (2010). PRISM: An introductory guide. Birmingham: Mott MacDonald/ Cambridge: RAND Europe.

Voorhees, A. M. (2013). A general theory of traffic movement. *Transportation*, *40*(6), 1105.

Waner, S. L. (1962). Stochastic choice of mode in urban travel. Evanston, IL: Northwestern University Press.

Williams, H. C. W. L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment and Planning, 9*, 284-344.

Williams, H. C. W. L. (1998). Congestion, traffic growth and transport investment: The influence of interactions and multiplier effects in related travel market. *Journal of Transport Economics and Policy, 32,* 141-163.

Williams, H. C. W. L., & Senior, M. L. (1977). Model-based transport policy assessment, 2: Removing fundamental inconsistencies from the models. *Traffic Engineering and Control, 18*, 464-469.

Wilson, A. G. (1967). A statistical theory of spatial distribution models. *Transportation Research, 1*(3), 253-269.

Wootton, H. J., & Pick, G. W. (1967). A model for trips generated by households. *Journal of Transport Economics and Policy, 1*, 137–153.

Appendix A. Random sampling method

Sometimes, the number of alternatives faced by individual is too large that estimating model parameters is very expensive or even impossible. With a logit model, estimation can be performed on a subset of alternatives without inducing inconsistency (Train 2009). For example, a researcher examining a choice situation that involves 100 alternatives can estimate on a subset of 10 alternatives, which have included the person's chosen alternative and 9 other alternatives randomly selected from the remaining 99 alternatives.

Suppose the full set of alternatives as F and a subset of alternatives as K, let q(K|i) be the probability under the researcher's selection method that subset K is selected given that the decision maker chose alternative i. We have q(K|i) = 0 for any K that does not include i. The probability that person n chooses alternative i from the full set is P_{ni}. The probability that person chooses alternative i conditional on the researcher selecting subset K is $P_n(i|K)$. So, the joint probability that the researcher selects subset K and the decision make chooses alternative i can be expressed as:

$$P_n(i|\mathrm{K}) = \frac{P_{ni}q(K|\mathrm{i})}{\sum_{j\in F} P_{nj}q(K|\mathrm{j})}$$

when we cancelled out the denominators of P_{ni}, above equation becomes:

$$P_n(i|\mathrm{K}) = \frac{e^{V_{ni}}q(K|\mathrm{i})}{\sum_{j\in F} e^{V_{ni}}q(K|\mathrm{j})}$$

and since the fact that q(K|j)=0 for any j not in K.

$$P_n(i|\mathrm{K}) = \frac{e^{V_{ni}}q(K|\mathrm{i})}{\sum_{j\in K} e^{V_{ni}}q(K|\mathrm{j})}$$

Suppose q(K|j) is the same for all $j \in K$, that means the researcher assigns an equal probability of selection to all nonchosen alternatives, so that the probability of selecting j into the subset when i is chosen by the decision make is the same as for selecting i into the subset when j is chose. McFadden (1978) calls this the "uniform conditioning property," since the subset of alternatives has a uniform probability of being selected conditional on any of its members being chosen by the decision make. When this property is satisfied, q(K|j) cancels out and above equation becomes the simply logit formula for a person who faces the alternatives in subset K.

The conditional log-likelihood function under this uniform conditioning property is:

$$CLL(\beta) = \sum_{n} \sum_{i \in K_N} y_{ni} \, ln \frac{e^{V_{ni}}}{\sum_{j \in K_n} e^{V_{ni}}}$$

Where K_N is the subset selected for person n. This function is the same as the log-likelihood function given in (x.x) except that the subset of alternatives K_n replaces, for each sampled person, the complete set. Maximization of CLL provides consistent estimators of β (Train 2009).

Appendix B. Example of a BIOGEME data file

[Choice]

// here the column containing the choice is defined

CHOICE

[ModelDescription]

// here the model can be described

[Beta]

// here the names, initial values, as well as the upper and lower

// limits of the model's parameters are defined

// also, it can be defined if the value is a fixed (1) or if it

// should be estimated (0)

// Name Value LowerBound UpperBound status (0=variable, 1=fixed)

B_TIME 0 -10 10 0

O_WOODLAND 0 -10 10 0

[Mu]

// In general, the value of mu must be fixed to 1. For testing purposes,

// you may change its value or let it be estimated.

// Value LowerBound UpperBound Status

1 0 1 1

[Utilities]

// in this location, the linear parameters of the model are defined

// Id Name Avail linear-in-parameter expression (beta1*x1 + beta2*x2 +

...)

```
1       A01_D01     D01_AV ASC_D01 * one + B_TIME * D01_TT_SCALED +
                    O_WOODLAND * D01_woodland
2       A02_D02     D02_AV ASC_D02 * one + B_TIME * D02_TT_SCALED +
                    O_WOODLAND * D02_woodland
3       A03_D03     D03_AV ASC_D03 * one + B_TIME * D03_TT_SCALED +
                    O_WOODLAND * D03_woodland
[Expressions]
// here are included those expressions which are not contained in the
// dataset
one = 1
D01_TT_SCALED = D01_TT / 60.0
D02_TT_SCALED = D02_TT / 60.0
D03_TT_SCALED = D03_TT / 60.0
//[Exclude]
// in case some variable needs to be excluded, here is the place to do it
//PURPOSE != 1
[Model]
// the type of model to be run is included in this section
$MNL // Multinomial Logit Model
```

Appendix C. Scenarios assumptions

This part is developed by Dr. Andrea Drewitt during her Ph.D.

Business As Usual 2044

Main aspects:

Additional to the mitigation proposals that are subject to planning of the Mersey Gateway, several projects are realised along the management objectives of respective management plans as part of the work carried out by the MGET.

The mitigation measures have been delivered according to the proposal of the biodiversity management plans.

Astmoor	No active management of saltmarsh vegetation. Implementation of saltmarsh restoration plan over the planning period, as per
Arpley	Development of country park after closure and capping of existing landfill after 2017. Subsequent impact on bird numbers, especially gulls which seem to be feeding on landfill, resting in the estuary's mudflats and surrounding areas (Hayley's dissertation).
Cuerdley	Remains of saltmarsh without major management due to landownership Reed bed management (cutting) is taking place on a seven year basis through the MGET.
Fiddlers Ferry	Will remain a site for electricity production throughout the planning period.
Gatewarth	Little active management due to lack of funding available. Natural succession throughout the planning period is expected.
Moore Nature Reserve	Less funding/management beginning with withdrawal of FCC as owner of the site due to closure of Arpley landfill. Long-term natural succession can be expected; temporary management works through the MGET likely.

Upper Moss Side	Continuous management of the site, but no priority management by Forestry Commission. Saltmarsh is grazed with appropriate cattle numbers. Visitor management is attempted.
Upper Moss Side Farm	Used as floodplain in the near future due to withdrawal of environment agency to maintain bunds. Long-term interest to manage site for biodiversity (estuary and farmland birds) i.e. maintaining agricultural land.
Oxmoor	Remains a local nature reserve.
Port Warrington	Further developed in the medium (>5 years). Used of Manchester Ship Canal will increase.
Spike Island	Will remain similar to present site.
St Helens Canal	No major change to the use of the canal/ Trans Pennine Trail will continue as a connection between Warrington and Widnes.
Tan House Lane	Partial development of brownfield site for mixed use (similar to planning permission submitted to Halton Borough Council, case no.: 05/00057/OUTEIA).
United Utilities	Management of site will continue as expected. Green areas mainly under natural succession (habitat without dog walkers/feral cats) (communication with Brian Tollitt), Himalayan balsam remains a problem.
Randles Island	No changes expected. Unknown.
Warrington Waterfront	Development for residential housing → connection has to be seen with the agricultural land as flood plains at Upper Moss Side Farm.
Widnes Warth	Saltmarsh restoration plan. Light grazing throughout the planning period.
Wigg Island	No change in management; small projects to be implemented by the MGET within the planning period.

Develop Boom 2044

Main aspects:

Less money for environmental purposes nation-wide

Green Belt land more likely to be under subject to exceptions to enable development

Other development made possible through change in legislation that simplifies the use of previously developed land.

Tidal barrage build downstream of the Upper Mersey Estuary as part of the development in the North West to produce energy from a reliable source to business and households.

Implementation of Lawton uncertain due continued lack of connectedness of natural areas.

Additional to the mitigation proposals that are subject to planning of the Mersey Gateway, small projects are realised along the management objectives of respective management plans as part of the work carried out by the MGET.

Astmoor	
Arpley	Development into country park, low maintenance design of park to avoid high cost.
Cuerdley	
Fiddlers Ferry	Continued energy production, potentially continued use of non-renewable energy production to support the grid in times of high demand.
Gatewarth	Natural succession of current habitats, no priority management due to lack of resources.
Moore Nature Reserve	Reduction of resources, therefore reduced management of the site and continued natural succession. Volunteer work likely to become important.
Upper Moss Side	Impacts of tidal barrage on saltmarsh. No priority site for forestry commission. Implementation of small projects through the support of the MGET.

Upper Moss Side Farm	Continued agricultural production. Use of site as potential flood plain.
Oxmoor	Low maintenance management through the overall situation of resource allocation.
Port Warrington	Full use of site, including an increased use of the Manchester Ship Canal.
Spike Island	
St Helens Canal	
Tan House Lane	Full development of site.
United Utilities	
Randles Island	Continued operation of the site. No change anticipated.
Warrington Waterfront	Developed for residential housing, flood defences installed to protect against potential flooding.
Widnes Warth	Implementation of the saltmarsh restoration plan according to planning permissions of the Mersey Gateway. Additional resources difficult to obtain for the maintenance of projects.
Wigg Island	

Natural Is Key 2044

Main aspects:

The UME is managed as a site that connects all the sites as one network. Exchange of information amount the stakeholders is crucial. Access to sites for management works and monitoring is possible. Data collected in the UME is stored in one central place.

Continued funding of projects over the planning period which enable management for biodiversity.

Additional to the mitigation proposals that are subject to planning of the Mersey Gateway, several projects are realised along the management objectives of respective management plans as part of the work carried out by the MGET.

Astmoor	Partially under management; implementation of the saltmarsh restoration plan throughout planning period.
Arpley	Country park development. Funding and initiative available to manage site for local residents and create a point to experience a view over large parts of the estuary; might include bird hides and environmental education.
Cuerdley	
Fiddlers Ferry	Continued energy production.
Gatewarth	
Moore Nature Reserve	Management of the reserve will continue, successful funding availability for long-term management.
Upper Moss Side	
Upper Moss Side Farm	Management of floodplains in conjunction with the neighbouring sites, e.g. for estuarine and farmland birds.
Oxmoor	

Port Warrington	Partial development of the site. Cooperation and participation in the nature work in the UME; work with the MGET (off-setting of construction works within the UME).
Spike Island	No major change to land use. Strengthened place within the UME as a visitor site. Communication of environmental actions possible there.
St Helens Canal	Work with Canal and River trust intensified; projects implemented to improve the site as a wildlife site. No use of the canal for commercial or private use for shipping.
Tan House Lane	No development of brownfield site for mixed use. Brownfield site under natural succession.
United Utilities	Management of site will continue as expected. Green areas mainly under natural succession (habitat without dog walkers/feral cats). MGET and UU are working together regarding monitoring of biodiversity (e.g. birds).
Randles Island	No changes expected. Unknown.
Warrington Waterfront	No development of site for housing, i.e. no change in green belt land. Land can be used as floodplain.
Widnes Warth	Continued light grazing for biodiversity benefits.
Wigg Island	Continued management through the MGET.